1960

Non-allelic gene interactions in a population of maize derived from a cross of two inbred lines

Angus Hillyard Hyer
Iowa State University

Mic 61–453

HYER, Angus Hillyard. NON–ALLELIC GENE INTERACTIONS IN A POPULATION OF MAIZE DERIVED FROM A CROSS OF TWO INBRED LINES.

Iowa State University of Science and Technology
Ph.D., 1960
Biology–Genetics

University Microfilms, Inc., Ann Arbor, Michigan

NON-ALLELIC GENE INTERACTIONS IN A POPULATION OF MAIZE DERIVED FROM A CROSS OF TWO INBRED LINES

by

Angus Hillyard Hyer

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major Subject: Crop Breeding

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State University
Of Science and Technology
Ames, Iowa

1960

TABLE OF CONTENTS

	Page
INTRODUCTION	1
LITERATURE REVIEW	2
MATERIALS AND METHODS	23
EXPERIMENTAL RESULTS	44
DISCUSSION	61
SUMMARY	68
LITERATURE CITED	71
ACKNOWLEDGMENTS	77
APPENDIX	78

INTRODUCTION

Genotypic variance can be divided into three components, additive variance, dominance variance, and non-allelic gene interaction or epistatic variance. The additive and dominance variance are associated with the effects of alleles at the same locus whereas non-allelic gene interaction variance arises from the interaction of genes at different loci. The nature of this interaction can be of many and varied forms, a few of which are considered in this thesis.

An evaluation of the relative importance of non-allelic gene interaction effects and an understanding of the nature of these gene actions are important to the development of the best breeding procedures to be used in a breeding program.

In this thesis a study was made of the presence or absence of non-allelic gene interaction effects in the inheritance of certain quantitative characters in a population of corn. An investigation was made also of the nature of the non-allelic gene interaction by comparing the fit of experimental data to that expected on the basis of four non-allelic gene interaction models.

The population considered in this study was obtained by crossing two homozygous lines and then selfing the resulting F_1. This leads to a population in which the frequency of the varying genes is one-half. The general procedure was to obtain measurements on quantitative characters in the resulting F_2, F_3, and F_4 generations and to make interpretations on the basis of these measurements.

LITERATURE REVIEW

Soon after the rediscovery in 1900 of Mendel's work, investigations led to the discovery that some characters were controlled by more than one factor pair and in many cases these independent factors were interacting with each other to produce an effect. Bateson et al. (1905) reported data on sweet pea that showed complementary gene action but it was not until the next year, Bateson et al. (1906), that a genetic explanation was given. Bateson and Punnett (1906) described an experiment in which a chicken with a pea comb was crossed to one with rose comb giving offspring with walnut comb. In the F_2 9 walnut : 3 pea : 3 rose : 1 single were obtained. They concluded that genes at two loci were interacting to produce the observed results.

Nilsson-Ehle (1909) gave the first explanation of inheritance of quantitative characters in Mendelian terms. Working with wheat kernel color, glume color and absence or presence of ligule in oats, he concluded that these characters were controlled by two or more independent factor pairs. From the results of this study he postulated that a continuous hereditary variation could appear in two ways; one by various combinations among several independent units, and second by interaction of these several independent units.

The terms "epistatic" and "hypostatic" were introduced by Bateson (1907) to describe a particular type of gene interaction. A gene which masked the effect of another gene at another locus was defined as being epistatic to the latter and the masked gene was referred to as hypostatic to the first.

Since the introduction of the term epistasis, this term has taken on new and different meaning to that proposed by Bateson (1907). Hollander (1955) reviewed the history of the use of this term in genetics. Fisher (1918) changed the spelling to epistacy and made it all inclusive by defining it in such a manner as to include all non-allelic gene interactions, this in a statistical sense. Wright (1935) used the term epistasis in reference to the interaction of genes at several loci for a model which was a special case of Fisher's formulation. Today the term epistasis is used in the literature in the same sense that Fisher defined epistacy.

In present day usage some authors use epistacy and others epistasis. In this thesis the term epistasis is used and is used synonymously with the term non-allelic gene interaction.

East (1913) in a study of inheritance of flower size in crosses between Nicotiana species concluded that factors affecting flower size were acting in a geometrical fashion and that the action was accelerative in that as each additional factor was added to the genotype the growth rate increased. Shull (1914) described a case in which the shape of the capsule in Bursa was controlled by duplicate factors. Although both Nilsson-Ehle (1909), in the case of red grain color in wheat, and East (1910), in the case of yellow endosperm in corn, had shown duplicate factor inheritance, Shull in his paper evidently was the first to refer to this type of gene action as duplicate factor inheritance.

Using a statistical approach to quantitative inheritance, Fisher (1918) partitioned genotypic variance into additive variance, dominance variance, and epistacy variance. He defined epistacy to be the deviation from additivity of effects between genes at two different loci. Genotypic

covariances between relatives in terms of dominance and additive variance were worked out for different relationships. He included a term for dual epistacy in some of the simpler covariances, dual epistacy being that associated with interaction between two loci. Although his results were not in a neat comprehensible form such as to exhibit the special cases as being particular forms of a general result, they were important in providing the basis for models including epistacy that were developed later by other workers.

Two models containing non-allelic gene interactions were proposed by Rasmusson (1934). These he called the logarithmic function model and geometric series model. After examining these models under various theoretical situations, he concluded that the geometric series model would fit biological data better than the logarithmic function model. In contrast to East's (1913) geometrical model, Rasmusson's geometrical model provided for diminishing effects as additional factors were added to a genotype. This would give a negative skewness to the F_2 distribution instead of a positive skewness as under East's hypothesis. Studying the time of flowering in Pisum, Rasmusson (1935) found that genes affecting this character were interacting. After studying diploid and tetraploid F_2 generations of a tomato species cross, Lindstrom (1935) attributed fruit size to additive, dominance, and geometrical gene action. He speculated that there were many types of genes affecting fruit size with many kinds of genetic interactions.

Neal (1935) grew single crosses, three-way crosses, and double crosses, and the F_2 of the double crosses of corn. By use of a formula developed by Wright (1922), he calculated the expected yield of the F_2. This formula

assumes additive effects with no epistatic effects. Neal obtained close agreement between the predicted and observed results which indicated epistatic effects were not important for yield in this material. Powers (1941) applying Neal's data to an additive model and a geometric model came to the same conclusion.

In a study of fruit size in Cucurbita pepo Sinnott (1937) concluded that this character was inherited geometrically. His conclusion was based on the positive skewness of the F_2 population. This would agree with East's (1913) hypothesis of geometric accelerative gene action.

Powers (1936) investigated the nature of the interaction of genes of four quantitative characters in a species cross of barley. He tested Rasmusson's (1934) theory of geometric inheritance and found that his data did not fit this model for any of the characters studied. His conclusion was that the nature of the non-allelic gene interaction followed no general pattern. By using genetic marker genes and their linkage relationships, Powers (1939) concluded that gene action was geometrically cumulative for the number of locules in a tomato hybrid he investigated. In another study on fruit weight of tomato Powers (1942) found his data fit a geometrical model better than an arithmetical one.

Evidence for geometrical gene action was found by MacArthur and Butler (1938) for fruit size in tomato, and by Charles and Smith (1939) for corolla tube length in tobacco. Khambanonda (1950) concluded that the gene effects for fruit weight in a red pepper cross were geometrical.

Kinman and Sprague (1945) made all possible single-cross combinations (45) between 10 inbred lines and produced the F_2 from these crosses. The inbreds, F_1's and F_2's were all grown in the same experiment. They cal-

culated expected yields based on an additive model and on a geometric model and compared the predicted with the observed values. Under both models they found significant differences between observed and expected although the data appeared to fit the additive model closer than it did the geometrical model. They suggested the deviations from the additive model may have been due to dominance or epistatic effects. Using the data of Kinman and Sprague, Jinks (1955) tested for epistatic effects by use of a diallel technique and found a significant estimate of epistasis.

A general formula for the mean of an inbred population was given by Kempthorne (1957, p. 444). This formula shows that one would get a non-linear relationship between the mean and the degree of inbreeding of an inbred population if there are non-allelic gene interactions of dominance x dominance, dominance x dominance x dominance, etc., deviations effects. One could get a linear relationship in the presence of epistatic effects if these involved only deviations with "additive" in their names, i.e., additive x additive, additive x dominance, etc., deviations. No epistasis would also give a linear relationship. Description of above terms like additive x additive, additive x dominance, dominance x dominance was given by Kempthorne (1954).

Stringfield (1950) from four homozygous inbred lines of corn made up a series of populations differing in amounts of heterzygosity. In studying three quantitative characters, yield, silking date, and ear height, he found a non-linear relationship between the performance for these characters and heterozygosity. As shown by Kempthorne (1957), this would indicate epistatic gene action was operative. In a similar study Sentz et al. (1954) also found a curvilinear relationship between heterozygosity and

performance in corn for yield, maturity and number of ears per plant. Field trials were conducted at four locations for four years so that their test covered a wide range of environments.

Smith (1952) grew in the same experiment four homozygous tobacco varieties, their six possible single crosses, and the three possible double crosses. He predicted the double cross performance for three quantitative characters by taking the average of the four non-parental single crosses. Finding no significant difference between predicted and observed values, he concluded this was evidence for the absence of epistatic effects. Such evidence is of doubtful validity because the prediction formula apparently includes some epistatic effects. He also thought that the absence of significant differences among the double cross means was further evidence for the absence of non-allelic gene interactions. The validity of such a test for detecting epistatic effects has not been established to the writer's knowledge. Using Mather's (1949) scaling tests to test for epistasis, Smith found no evidence for such gene action in the characters he studied.

Mather and Vines (1952) found evidence for non-allelic gene interactions in a tobacco cross for the characters plant height and time of flowering. They used F_1, F_2, F_3, F_4, and backcross generations and biparental progenies from the F_2 generation. The material was grown for three years. They used a model that included additive, dominance, and environmental effects. Expected values were calculated based on this model. The sum of squares of the deviation of expected from observed was divided into three parts, one of which included epistatic effects. By appropriate statistical methods they were able to test the significance of

this term. Evidence of interaction between epistasis and years was found for the plant height character but not for flowering time. The failure to find a scale effectively transforming the data to an additive scheme was given as further evidence for the presence of epistasis.

Opsahl (1956) reported evidence of epistatic effects in tobacco for flowering time but not for plant height in material derived from a cross of the same two homozygous parents used by Mather and Vines (1952). Opsahl grew his material in a different year and location than did Mather and Vines. The failure of Opsahl to detect non-allelic gene interaction effects for plant height may be a manifestation of the epistatic-environmental interaction Mather and Vines found for that character.

In estimating additive and dominance variance of several characters in tobacco by use of biparental progenies, Robinson et al. (1954) obtained estimates indicating little dominance variance. Their models did not include a term for epistatic effects, therefore these effects would tend to bias upward the dominance variance estimates. They concluded the low magnitude of the dominance variance indicated little effect from epistasis.

Powers (1955) analyzed yield data of a tomato hybrid by five methods: (a) analysis of variance of means, (b) partitioning the variance into genetic and environmental components, (c) dividing weight per fruit into component characters, (d) calculating the relative percentages of variances accounted for by regression, and (e) partitioning the frequency distributions of the segregating populations on the basis of certain genotypes. By these methods he found evidence of epistatic effects and was able to postulate the nature of some of the non-allelic gene interactions. He concluded that no one method provided all the information available

but they were supplemental to each other.

Expected values of sample covariances and variance components in terms of genotypic covariances and variances were derived by Horner and Weber (1956) for populations produced by crossing two homozygous lines and subsequent self fertilization. Using a completely additive model, an additive model with dominance, and an additive model with dominance and additive x additive interaction, they compared expected and observed covariance and variance values for the character maturity in populations of F_2 to F_7 derived from a cross of two homozygous lines of soybeans. They found that the completely additive model fitted the data as well as the other two models and that 96 percent of the variation among the sample covariances and variance components could be explained by the completely additive model.

In a study of three open-pollinated varieties of corn and their F_1, F_2, and backcross generations Pollak et al. (1957) concluded that epistatic effects for yield, although present, were not of sufficient magnitude to be of importance.

Gamble (1957) studied data obtained from six homozygous inbred lines of corn and their F_1, F_2, and backcross generations at two locations for two years. Data on six quantitative characters, plant height, number of kernel rows, ear length, ear diameter, weight per 100 seeds, and yield, were obtained and the data analyzed by a method developed by Anderson and Kempthorne (1954) to detect non-allelic gene interactions. He found that both nonepistatic and epistatic gene action were important in the inheritance of the characters studied, but that nonepistatic gene action was relatively more important than was epistatic effects. Nonepistatic effects

were little influenced by the environment, whereas epistatic gene action appeared to be interacting with environment. Epistasis was slightly less important for yield than it was for the other characters studied. Some lines contributed more epistatic gene action to the inheritance of the characters investigated than did others. Gamble also found a relationship between epistasis and heterosis. In general, a cross exhibiting a high degree of heterosis also showed epistatic gene action.

Heterosis of a cross between two homozygous parents was expressed in terms of additive, dominance, and non-allelic gene interaction effects by Jinks and Morley Jones (1958). Heterosis was defined as $\bar{F}_1 - \bar{P}_o$, where $\bar{P}_o$ is the mean of the better parent. Hayman (1958b) found that the additive effects of Jinks and Morley Jones were not free of epistatic effects as the additive effects included a term r, which Jinks and Morley Jones called a measure of the degree of association of loci. Jinks and Morley Jones applied their model to plant height data of several tobacco crosses. They concluded that in these crosses "The presence or absence of heterosis is not in itself indicative of the presence or absence of any particular type of gene action or interaction; it can result from a whole range of combinations of gene effects, . . ." Nevertheless they found a positive correlation between the presence or absence of heterosis and the presence or absence of epistasis. Heterosis occurred with greater frequency in the presence of epistasis and at a higher level of expression than it did in the absence of epistasis.

Robinson et al. (1958) used two open-pollinated varieties of corn and the cross between them in a study of genetic variance components for yield. Their observed results were in close agreement with those expected based

on an additive model with partial to complete dominance. The authors stated that the results offered no evidence for or against the presence of epistasis because the data may not have been able to distinguish between an additive model and one including epistasis. Such a situation was shown by Horner et al. (1955).

Depression of means with inbreeding was studied by Ryder (1958) in two strains of lima beans and their F_1, F_2, F_3, and backcross generations for the character seed size. He constructed three types of complementary gene models, additive complementary, dominant complementary, and geometric complementary. These were based on two loci with two alleles per locus. His data fitted all three models equally well. Scaling tests and skewness of the generation frequency distribution and means also gave evidence of non-allelic gene interaction. Ryder concluded that a form of complementary gene action contributed to the inheritance of seed size.

Bauman (1959) made comparisons between single crosses and three-way crosses of corn. The single crosses were made up by crossing two lines, say A and B, to a common inbred, say C, and the three-way cross was produced by using A x B as the single cross and C as the inbred parent. The performance of the three-way cross was then predicted on the basis of 1/2 [(A x C) + (B x C)] i.e. the average of the two single crosses. He compared predicted and observed values for yield, ear height, and kernel row number. Wherever a significant difference was found between observed and predicted, this was taken as evidence of non-allelic gene interaction effects. He was able to make sixteen such comparisons for each character. The material was grown for three years at one location. He found several significant deviations among all characters for single year data but no

such differences in the combined data over years. The power of the test for the combined analysis was low which may account for the non-significant results. Significant interaction between epistasis and years was reported. In several cases two different testers (C lines) were used with the A and B lines. Bauman found evidence that the tester genotype influenced the detection of epistasis.

The method used by Bauman to detect epistasis could be criticized on the same basis as that used in questioning the prediction formula used by Smith (1952). The two prediction formulas use the same principle in that they use the non-parental single crosses to predict the double or three-way cross. As pointed out previously, such prediction formulas apparently include some epistatic effects.

Barley data of fifteen crosses made up from six varieties gave a good fit to a geometrical model according to Grafius (1959). The geometrical model assumed yield is the volume of a rectangular parallelepiped with edges X, Y, and Z equal to heads per plant, seeds per head, and average seed weight respectively. Grafius concluded that the F_1 vigor for yield was due to epistasis.

Fisher et al. (1932) and Mather (1949) suggested that the problem of non-allelic gene interactions could be circumvented by the choice of a proper measurement scale which would make the effects of genotypes at different loci additive. Horner et al. (1955) discussed this problem and concluded that scale transformation probably would not materially lessen the problem of non-allelic gene interactions. They suggested that in order to handle the problem of non-allelic gene interactions, models should include such terms. Cockerham (1959) stated that transformations may be useful

for particular types of gene action, but doubted their usefulness in real situations where many types of gene action probably are involved.

In recent years models have been derived that have included epistatic terms. Kempthorne (1954) partitioned the genotypic variance into additive, dominance, and epistatic variances. He further subdivided the epistatic variance into components arising from interactions among sets of 2, 3, 4, ... n loci and showed that the interaction variance is composed of portions attributable to the interaction of additive and/or dominance effects of genes. His model was general for arbitrary number of alleles and loci with the assumption of no linkage. Covariances between relatives in a random mating population were expressed in terms of the genotypic variance components. These results were given in a form more suitable for the user of genetic statistics by Kempthorne (1955).

Anderson and Kempthorne (1954) developed a model which included epistatic terms. They called their model a factorial gene model because it was an adaptation of the factorial model used in experimental designs. Their model included no linkage, gene frequency equal to one half and no multiple alleles. By the use of this model a general formula for the mean genotypic value of a population derived by n generations of selfing was obtained. This formula included terms involving epistatic effects. They applied the model to red pepper data of Khambanonda (1948). These data are also reported by Khambanonda (1950). Anderson and Kempthorne concluded epistatic effects may have been important in Khambanonda's material. They also detected epistatic effects in some corn data they analyzed. Applying the model to the scaling tests of Mather (1949) they found that these tests are not exact for detecting epistasis as Mather's tests include

three-factor and higher-factor interactions.

Cockerham (1954), also using a factorial scheme, developed models including non-allelic gene interactions. The hereditary variance of diploid populations having no multiple alleles was partitioned into additive variance, dominance variance, and variances containing epistatic effects. Covariances between hereditary values of relatives were partitioned into similar variance components.

Horner et al. (1955) derived expectations of genotypic means, variances, and covariances for the following non-allelic gene interaction models: complementary, duplicate factor, multiplicative, and optimum number. Assumptions common to all models were: no multiple alleles, no linkage, no selection, and symmetry of genotypic effects. Symmetry was defined as that condition in which "all genotypes having the same number of (++), (+-), and (--) loci, respectively, have equal value or effect." They used these models to evaluate the bias due to epistasis in estimating average degree of dominance, additive variance, and dominance variance in the three experiments described by Comstock and Robinson (1952) wherein they assumed no epistasis. In several cases the bias due to epistasis was found to be serious.

Horner et al. (1955) showed when epistasis is present that estimating average degree of dominance under certain situations by the methods of Comstock and Robinson (1952) would give an estimate of overdominance when in fact dominance was only partial or complete. Using the four models, the bias due to non-allelic gene interaction effects when estimating additive variance by doubling the parent-offspring regression also was found to be serious under certain conditions. They proposed a method for de-

tecting epistasis in a population derived by self fertilization following the crossing of two homozygous lines. The failure of population means in the presence of epistasis to be a linear function of heterozygosity also was examined.

Kempthorne's (1954) results were applied to the complementary, duplicate factor, multiplicative, and optimum number models of Horner et al. (1955) by Horner and Kempthorne (1955). They considered the case of a random mating population, gene frequency one half, two alleles per locus, symmetry of genotypic effects, and no linkage. They found that under several of the conditions that are likely to be encountered under actual genetic situations that it would be hard to distinguish between an additive model with dominance and a multiplicative model when gene frequencies are one half. They considered the biases under these models in parent-offspring and full-sib covariances when they are used as estimates of additive variances. The results of Horner and Kempthorne were extended by Horner (1956a, 1956b) to the case in which the frequency of the allele with the greater effect is the same at all loci, not necessarily equal to one half.

Kempthorne (1956b) derived the expected covariances between relatives in diploid populations under selfing for the case of an arbitrary number of alleles, arbitrary number of loci, and arbitrary epistasis, in the absence of linkage and selection.

Hayman (1958b) gave a model with epistatic terms and described a method of testing for epistatic effects by the use of two homozygous lines and their F_1, F_2, and back-cross generations.

The variances of the models of Horner et al. (1955) were partitioned

into additive, dominance, additive x additive, additive x dominance, dominance x dominance, and higher-order variances by Cockerham (1959). He examined the effects of limiting the genotypic coefficient of variation for the multiplicative model, as was done by Horner et al. They limited the coefficient of variation to 30 percent. Cockerham concluded that if the genotypic coefficient of variation was limited to 40 percent or less it would be essentially impossible to distinguish between an additive model and a multiplicative one. The complementary and duplicate factor models were shown to be variants of the multiplicative model. Cockerham stated that the feasibility of distinguishing between different gene models is questionable on three counts: "(1) gene action for biometrical characters is probably not all of one type (2) variants of different models give similar results and (3) estimates from data generally have too large sampling errors to be very discriminatory."

Within the last decade a considerable number of papers have appeared concerning the use of the diallel table in quantitative genetic studies. Griffing (1950) discussed the theoretical aspects of certain statistics that can be calculated from a diallel table in which homozygous parents are used. The main statistic considered was the constant parent regression. An approach to the study of quantitative inheritance was shown which consisted mainly in constructing different gene models and choosing the one that best fits the data as determined by various statistics. Hayman (1954a, 1954b) considered the theory of diallel crosses between homozygous lines in terms of additive and dominance effects. Although his models did not include a term for non-allelic gene interaction effects, he briefly discussed the effects of such terms on his models.

A method of testing for epistatic effects by the use of diallel data, where the parents are homozygous, was developed by Jinks (1954). This test is based on what Jinks designated as the regression of array covariance on array variance. Using this method of detecting epistatic effects, Jinks (1955) analyzed published data of a number of diallel crosses and related crossing schemes of maize, flax, egg plants, Galeopsis species, and tobacco. By his method he detected considerable epistasis in these species for various quantitative characters. He also calculated estimates of the average degree of dominance. Whenever he obtained an estimate of overdominance, he also obtained a significant non-allelic gene interaction. Reanalyzing the data, after omitting all crosses showing significant non-allelic gene interaction, Jinks found that in all cases a decrease in the degree of overdominance was obtained. In one case, in tobacco, overdominance completely disappeared to an estimate of only complete dominance. Using the data on corn of Kinman and Sprague (1945) he found that the mean yield of F_1 families showing epistasis was 90 bushels per acre whereas those showing no epistasis had a mean yield only of 77 bushels per acre, a difference of 13 bushels.

Jinks (1956) extended his 1954 analysis of diallel crosses to include the F_2 and backcross generations. Using the segregating generations along with the F_1 generation provided three methods of detecting epistasis: (a) regression of array covariance on array variance, (b) scaling tests, and (c) homogeneity of the least squares estimates of the components of variation over statistics. The third method is the technique of Mather and Vines (1952), wherein they divided their sum of squares of deviations of expected from observed into three parts. Jinks applied these three methods

to tobacco data studying the characters plant height and flowering time. All tests detected epistatic effects for plant height. Epistatic effects were detected for flowering time but the regression of array covariance on array variance method appeared to be the least efficient of the methods in detecting epistasis for this character. From the shape of the regression curve of array covariance on array variance, he concluded the type of gene action involved in the inheritance of the plant height character was of a complementary type and the type involved for the flowering time character was a duplicate factor type. His conclusions were that duplicate factor gene action may require method 2 or 3 to be detected thus requiring the F_2 and backcross generations to be grown whereas F_1 data will detect complementary gene action by use of method 1.

Dickinson and Jinks (1956) extended the methods of Jinks (1954) and Hayman (1954b) to the general case in which the parents of the diallel crosses may be homozygous or heterozygous.

Using the methods developed by Jinks (1954) and Hayman (1954a, 1954b), Allard (1956a) analyzed a 9 x 9 diallel of lima beans and reported evidence that complementary gene action was important in the inheritance of seed size. Allard (1956b) extended the methods of Jinks (1954) and Hayman (1954a, 1954b) to one in which an assessment of the interaction of additive, dominance, and epistatic effects with environment is possible. He applied this method to a diallel situation involving intervarietal hybrids of Nicotiana rustica. The material was grown for three years at one location. Allard found epistasis to be unimportant for flowering date but important for plant height. The epistatic effect for plant height was suggested to be of a complementary type and appeared not to be interacting

with environment.

A general approach to the genetical theory and analysis of the diallel table was considered by Kempthorne (1956a). He developed an analysis which included multiple alleles, arbitrary number of loci, and arbitrary epistasis. Six quantities of genetical interest which can be estimated by use of the general theory were considered. Kempthorne questioned the logical value of removing some of the parents from the diallel table and considering the remainder, as was done by Jinks (1955).

Since Sprague and Tatum (1942) defined general and specific combining ability, these terms have appeared frequently in the literature. Sprague and Tatum considered that general combining ability contained mainly additive effects and specific combining ability mainly dominance and epistatic effects. Griffing (1956) showed that for the general case of multiple alleles, arbitrary number of loci, and homozygous parents that general combining ability variance includes additive variance plus part of the additive interaction variances and that specific combining ability variance contains dominance variance plus the additive and/or dominance interaction variances. Matzinger (1956), by using a model of only two loci and multiple alleles, arrived at essentially the same results as did Griffing.

Matzinger and Kempthorne (1956) extended the results of Matzinger (1956) to include an arbitrary number of loci and to the case in which the parents are all of the same degree of inbreeding, not necessarily completely homozygous. They also gave the expected mean squares of diallel crosses repeated over locations and years in terms of genotypic variances and interactions of genotypic components with environment.

Hayman (1957) extended Mather's (1949) C test for epistasis to a diallel cross situation. A method of distinguishing between duplicate factor and complementary gene action was given. He discussed methods of investigating the relationship between epistasis and average degree of dominance and between epistasis and heterosis. After examining published diallel data of tobacco, corn, and cotton by these methods, Hayman concluded that "Heterosis is a composite phenomena: possible causes are epistasis, overdominance and accumulation of favorable dominants in the heterozygotes."

The diallel analysis of Hayman (1954b) was extended by Hayman (1958a) to include F_2 families. He found that this greatly increases the accuracy of measurement of the components of genetic variation and makes it possible to distinguish some of the more complicated genetic systems that might be present. Jinks and Stevens (1959) extended the results of Jinks and Morley Jones (1958) to include the diallel set as a whole and not just the individual crosses.

Diallel theory was applied to an 8 x 8 diallel table of corn data by Rumbaugh and Lonnquist (1959). They found little evidence of epistatic effects in the inheritance of yield, plant height and ear height.

Most models of quantitative inheritance considered to date include the assumption of no linkage effects. In actual genetic situations it is known that this assumption is not valid and it is recognized that epistatic effects probably are biased by linkage relationships. Very little work has been done on this problem. Mather and Vines (1952), by using segregating generations and growing duplicate plots, were able to separate the sum of squares of the deviations of expected from observed results into

three components (1) linkage effects (2) residual interaction effects, and (3) duplicates error. The residual interaction effects includes epistatic effects so by this method linkage effects and epistatic effects are separated.

Hayman and Mather (1955) considered models containing two-factor interactions and gave a method of separating these two-factor interactions from linkage effects by the use of second degree statistics of F_2 and F_3 generations. Following the methods of Hayman and Mather, Opsahl (1956) developed a method of discriminating between non-allelic gene interaction effects and linkage effects by the use of parents, F_1, F_2, and backcrosses of the F_2 to the F_1 plants and to the parents. This model included only two-factor interactions. Opsahl applied this method to tobacco data but found that the data were not sufficient to discriminate between linkage and epistatic effects.

The effects of linkage on first-order statistics were discussed briefly by Kempthorne (1957, p. 483), and on second-order statistics of populations derived from a pair of inbred lines by Mather (1949).

The effects of linkage on the theoretical covariances between relatives in a random mating population, whose genotypic frequencies were in linkage equilibrium, were considered by Cockerham (1956). He found that if one relative is the ancestor of the other that covariances are not affected by linkages if position effects are absent, but if one relative is not the ancestor of the other, the covariances are affected by linkage and the bias appears in the epistatic components and not in the additive or dominance components.

Jinks and Stevens (1959) developed a technique of separating epistatic

effects from another term called correlated gene distribution by the use of diallel data. Correlated gene distribution was defined as a measure of the distribution of dominant and recessive alleles in the parents so would include linkage effects. Applying this technique to tobacco data of Jinks (1954, 1956) they found that the correlated gene distribution was not an important source of variation among the interacting genes, and therefore concluded that linkage was not biasing the estimates of epistasis in this material to any great degree.

MATERIALS AND METHODS

Materials and Experimental Procedure

The material used in this study was derived from a single cross of two inbred lines of Zea maize, Hy and B10. The cross was made in 1949. The two inbred lines were assumed to be homozygous at all loci. The line B10 had been selfed for eight generations. The exact number of selfing generations for Hy was not known but probably was well over fifteen. This line was one of the lines developed in the early 1930's and had been maintained by selfing since that time.

In 1951 the F_2 generation of the Hy x B10 cross was grown at Ames, Iowa and single ears of 90 random plants were selfed. At harvest time the 90 selfed ears were harvested, dried to a uniform moisture, and then ear measurements taken.

The seed from the 90 selfed F_2 plants were stored until 1954 at which time the F_3 generation was grown at Ames, Iowa. The F_3 was grown in F_2 progeny rows of 25 plants. The plants were spaced approximately 13 inches apart within the row and the rows were 40 inches apart. Each F_2 progeny row was replicated twice and the F_3 material grown in a randomized block design. Single ears of four plants were selfed within each row; thus within each F_2 sub-population in the F_3 generation a total of eight F_3 plants were selfed.

An F_2 sub-population in the F_{2+k} generation is defined as consisting of all the F_{2+k} plants that are descendants of the same F_2 plant; an F_3 sub-population consists of all the F_{3+k} plants that are descendants of the same F_3 plants, etc., where $k \geq 0$. Where $k = 0$ the F_n plants themselves

are the sub-population, e.g. an F_2 sub-population in the F_2 generation would be an F_2 plant itself. This terminology is used for consistency, since reference will be made to individual plants within an F_n generation as F_n sub-populations in the F_n generation.

In most cases all eight selfed F_3 plants within an F_2 sub-population produced an ear; but in some instances less than eight were available for harvesting. In the case of three plots (rows) no ears were produced. Missing plot values were calculated for these plots. In the fall the selfed ears were harvested and ear measurements taken after drying the ears to a uniform moisture content.

Of the 90 F_2 sub-populations grown in the F_3 generation, only the 64 that were grown in the F_4 generation were analyzed by the analysis of variance technique.

In 1957 the F_4 generation consisting of F_3 progeny rows was grown at Ankeny, Iowa. Only 64 F_2 sub-populations were used in this experiment. These 64 were chosen at random from the 90 available. Of the eight or less F_3 sub-populations within each F_2 sub-population available in the F_4 generation, only four were grown. These four were chosen at random from those available, except for the restriction that a minimum of 120 seeds must have been produced by the F_3 plant to provide enough seed for planting purposes. Thus a total of $4 \times 64 = 256$ entries were used for the 1957 experiment.

Because of the large number of entries, the experiment was separated into eight sub-groups called <u>sets</u> to reduce the size of the replication. Each set consisted of eight random F_2 sub-populations and four F_3 sub-populations within each F_2 sub-population or a total of 32 entries. Each

set was grown in a randomized block design of six replications. Each set was analyzed separately, and then the eight sets combined into one analysis.

Plot size was one row of ten plants with the plants spaced approximately 13 inches apart in the row. To insure a good stand the plots were seeded to two seeds per hill and after emergence plants thinned to one per hill. Open pollination was allowed in the experiment.

From each plot single ears from five plants, if available, were harvested, dried to a uniform moisture level, and then measurements taken. Where possible, only those plants which were guarded by other plants in the row were harvested. Because of poor pollination and ear set, this was not always possible and in some cases unguarded plants were harvested. No record was kept as to which harvested plants were guarded and which were not. Some plots had less than five plants that produced ears and a few had no plants producing ears. Missing plot values were calculated for those plots having plants with no ears. In the expected mean squares of the analysis of variance the harmonic mean of plants per plot was used to adjust for unequal numbers of plants per plot.

The ear measurements taken on the F_2, F_3, and F_4 material were number of kernel rows, ear length, ear diameter, and total shelled seed weight. Ear diameter was measured at the thickest part of the ear. These characters are known to be quantitatively inherited.

Statistical Procedures

Horner et al. (1955) developed a method of detecting non-allelic gene interactions using genetic variances and covariances. This method is used

in this study as a test for non-allelic gene interactions. The method will be described in what follows with some changes in notation from that of Horner et al.

For the case of one locus with two alleles genotypic values may be represented by the following:

Phase at i^{th} locus	Frequency	Genotypic value	Coded genotypic value
(++)	q_i^2	$z_i + 2u_i$	u_i
(+-)	$2q_i(1-q_i)^2$	$z_i + u_i + a_iu_i$	a_iu_i
(--)	$(1-q_i)^2$	z_i	$-u_i$

Here the genetic state at a given locus is referred to as the phase at the locus. One of the two alleles at a locus is called by convention the plus (+) allele and the other the minus (-) allele, though no significance is to be attached to these words, since they appear symmetrically and which is the better allele is determined by the quantity u_i defined below. For a two allele diploid case the possible phases at a locus are the homozygous plus (++), the heterozygous (+-), and the homozygous minus (--) phases.

The quantities q_i, u_i, a_iu_i, and a_i are defined as:

q_i is the relative frequency of the plus gene at locus i.

u_i is half the difference between the mean of the values of all genotypes having the (++) phase at locus i and the mean of all genotypes having the (--) phase at locus i.

a_iu_i is the genotypic value of the heterozygote minus the mean of the corresponding homozygotes with respect to the i^{th} locus.

a_i is a measure of dominance and is interpreted as follows:

$a_i = 0$	No dominance
$0 < a_i < 1$	Partial dominance of plus allele
$-1 < a_i < 0$	Partial dominance of minus allele
$a_i = \pm 1$	Complete dominance
$a_i > 1$ or $a_i < -1$	Over dominance

Several different genetic models were considered by Horner et al. (1955). The assumptions common to these models were:

1. Normal diploid behavior at meiosis
2. No multiple alleles
3. No linkage
4. No selection
5. Symmetry

Symmetry was defined as the condition in which "all genotypes having the same number of (++), (+-), and (--) loci, respectively, have equal value or effect." In the general case of s interacting loci this means that

$$u_1 = u_2 = \dots u_s$$

$$a_1u_1 = a_2u_2 = \dots a_su_s$$

$$\sigma_A^2 = \sum_i \sigma_{A_i}^2 = s\,\sigma_{A_i}^2$$

$$\sigma_D^2 = \sum_i \sigma_{D_i}^2 = s\,\sigma_{D_i}^2$$

where σ_A^2 is the total additive variance and σ_D^2 the total dominance variance, and $\sigma_{A_i}^2$ and $\sigma_{D_i}^2$ are the same quantities associated with the i^{th} locus. This symbolism for additive variance and dominance variance will be used throughout this thesis.

The value of a genotype of an individual may be represented by

$$Y_{y_0,y_1,y_2}$$

where

y_0 = number of (--) loci

y_1 = number of (+-) loci

y_2 = number of (++) loci

and

$$y_0 + y_1 + y_2 = s = \text{total number of interacting loci.}$$

The genetic models considered and their equations are

1. Completely additive model

$$Y_{y_0,y_1,y_2} = (y_1 + 2y_2)u$$

Here the genotypic value is proportional to the number of plus genes in the genotype since $y_1 + 2y_2$ is the number of plus genes.

2. Additive model with dominance

$$Y_{y_0,y_1,y_2} = (y_1 + 2y_2)u + y_1 au$$

This differs from the completely additive model in that each heterozygous locus makes a contribution au to the value of the genotype.

3. Complementary model

$$Y_{y_0,y_1,y_2} = 0^{y_0} = \begin{cases} 0 \text{ when } y_0 > 0 \\ 1 \text{ when } y_0 = 0 \end{cases}$$

No effect is produced if there are any (--) loci in the genotype.

4. Duplicate factor model

$$Y_{y_0,y_1,y_2} = 1 - 0^{y_1 + y_2} = \begin{cases} 0 \text{ when } y_1 + y_2 = 0 \\ 1 \text{ when } y_1 + y_2 > 0 \end{cases}$$

An effect is produced if there are any plus alleles in the genotype. No effect is produced when all interacting loci are at the (--) phase.

5. Multiplicative model

$$Y_{y_0,y_1,y_2} = e^{y_1}b^{y_2}$$

A locus in the (+−) phase increases the genotypic value by the factor e and a locus in the (++) phase by the factor b. When e = b there is complete dominance. The case of e > 1, $b = \sqrt{e}$ corresponds to no dominance on the logarithmic scale. Only the case of complete dominance for this model is investigated in this study.

Horner et al. (1955) gave evidence to show that the range b = 1.1 to b = 1.4 is a reasonable one to represent actual conditions. Later Horner and Kempthorne (1955) showed b = 1.06 could be considered as a reasonable lower limit. Therefore, in this study the range from b = 1.06 to b = 1.4 is included for investigation.

This model may be changed to an nonepistatic one by a logarithmic transformation and for this reason is not considered as true non-allelic gene interaction by many workers. It is included here as a non-allelic gene interaction model because of the practical problem of not knowing when and how to transform data. This point is discussed more fully by Horner et al. (1955).

6. Optimum number model

$$Y_{y_0,y_1,y_2} = -[y_1 + 2y_2 - (s + d)]^2$$

This model represents a special case of a more general one considered by Wright (1935). The quantity (s + d) is considered as an optimum number of plus genes because the value of the genotype is greatest where $y_1 + 2y_2 = s + d$. Since the possible number of plus genes, $y_1 + 2y_2$, may be between 0 and 2s, d may take any value between -s and s.

Models (3), (4), (5), and (6) are models involving non-allelic gene interaction. To detect these types of gene actions from that of complete additivity or additivity with dominance, use is made of covariances and variances of relatives.

If an individual X_i in generation k-1 is self-pollinated, m offspring in generation k will be obtained. Continued selfing these m offspring to generation n and n', where $2 \leq k \leq n \leq n'$, will result in m F_k sub-

populations in the F_n and $F_{n'}$ generations. An F_k sub-population in the F_n generation is, as defined previously, all the individuals in the F_n generation that are descendants of the same F_k individual. We may calculate a covariance between the m means of the F_k sub-populations in generation n and n'. Similar covariances for each individual in generation k-1 may be calculated. If each of these covariances is weighted by the frequency of respective X_i in generation k-1, we will obtain another covariance which is designated as $Cov(X,X'/k,k-1;n,n')$ and is defined as the covariance within F_{k-1} sub-populations of F_k sub-population means in the F_n and $F_{n'}$ generations. If $n = n'$, $Cov(X,X'/k,k-1;n,n')$ becomes a variance but for consistency of symbolism it is left in the above form.

Making an additional assumption to the five previously listed that gene frequency at all loci is equal to one half and using properties of the multinomial distribution, the expected value of $Cov(X,X'/k,k-1;n,n')$ for the various models can be shown to be as follows, Horner et al. (1955):

Model	$Cov(X,X'/k,k-1;n,n')$
Completely additive	$(\frac{1}{2^{k-2}})\ \sigma_A^2$
Additive with dominance	$(\frac{1}{2^{k-2}})\ \sigma_A^2 + \frac{1}{2^{n+n'-k-2}}\ \sigma_D^2$
Complementary	$(G)^s - (G')^s$
Duplicate factor	$(H)^s - (H')^s$
Multiplicative (b = e)	$(b^2G + bL + H)^s - (b^2G' + bL' + H')^s$
Optimum number	$\frac{sd^2}{2^{k-3}} + \frac{s}{2^{n+n'-k}} + \frac{2^k - 3}{2^{2k-3}}\ s(s-1)$

where $2 \leq k \leq n \leq n'$

$$\sigma_A^2 = \frac{su^2}{2}$$

$$\sigma_D^2 = \frac{sa^2u^2}{4}$$

$$G = \frac{2^k + 2^n + 2^{n'} + 2^{n+n'} - 2^{n+n'-k}}{2^{n+n'+1}}$$

$$G' = \frac{2^{k-1} + 2^n + 2^{n'} + 2^{n+n'} - 2^{n+n'-k+1}}{2^{n+n'+1}}$$

$$H = \frac{2^k - 2^n - 2^{n'} + 2^{n+n'} - 2^{n+n'-k}}{2^{n+n'+1}}$$

$$H' = \frac{2^{k-1} - 2^n - 2^{n'} + 2^{n+n'} - 2^{n+n'-k+1}}{2^{n+n'+1}}$$

$$L = \frac{2^{n+n'-k} - 2^k}{2^{n+n'}}$$

$$L' = \frac{2^{n+n'-k+1} - 2^{k-1}}{2^{n+n'}}$$

In the case of the non-allelic gene interaction models

$$\mathrm{Cov}(X,X'/k,k-1;n,n') = \frac{\sigma_A^2}{2^{k-2}} + \frac{\sigma_D^2}{2^{n+n'-k-2}} + \Delta(i)$$

where $\Delta(i)$ is a function of the non-allelic gene interaction effects. Now if both sides of the above equation are multiplied by 2^{k-2} we obtain

$$2^{k-2}\,\mathrm{Cov}(X,X'/k,k-1;n,n') = \sigma_A^2 + \frac{\sigma_D^2}{2^{n+n'-2k}} + 2^{k-2}\,\Delta(i)\ .$$

The quantity $2^{k-2}\,\mathrm{Cov}(X,X'/k,k-1;n,n')$ is defined as the adjusted Cov $(X,X'/k,k-1;n,n')$. If all genetic variance is additive, the adjusted $\mathrm{Cov}(X,X'/k,k-1;n,n')$ is equal to σ_A^2.

The ratio of two adjusted covariances will equal one if all genetic variance is additive variance, but will not necessarily be one if the genetic variance includes dominance variance and/or non-allelic gene interaction variance.

For the additive model with dominance the ratio of two adjusted covariances will equal one if the quantity $(n+n'-2k)$ is equivalent for the two covariances. If this quantity is not equivalent, then the ratio will be something other than one.

A technique for detecting non-allelic gene interaction effects, as suggested by Horner et al. (1955) and used in this study, is to compare estimated ratios of adjusted covariances with expected ratios, and if the estimated ratios differ significantly from one and also from that expected based on the additive model with dominance, then this is considered as evidence for the presence of non-allelic gene interaction effects.

This technique may not detect a particular type of non-allelic gene interaction when a particular ratio is used because the expected ratio based on the epistatic model will not be different enough from the expected ratios based on the nonepistatic models to be detected by experimental data. Horner et al. showed that certain ratios were better than others in distinguishing between epistatic effects and nonepistatic effects and that certain ratios would be good to detect a particular type of non-allelic gene interaction effect if it existed, whereas other ratios would be required to detect other types of epistatic effects. By the proper choice of ratios one may be fairly sure of detecting a particular type of non-allelic gene interaction if it is present.

Upon finding evidence for non-allelic gene interaction effects, one then can try to determine the nature of the gene action. The procedure would be to compare the estimated ratio of adjusted covariances to that expected based on epistatic models. In this study the four epistatic models listed above were used for such comparisons.

The four epistatic models investigated here do not exhaust the possible types of non-allelic gene interactions, therefore the nature of the gene action, if epistasis is present, may not be determined by considering only these four models.

Using the F_2, F_3, and F_4 data obtained in this investigation the following seven covariances and variances can be estimated. Also shown is the general method by which each covariance or variance was estimated.

Covariance	How estimated
Cov(X,X'/2,1;3,3)	Analysis of variance of F_3 data
Cov(X,X'/2,1;2,3)	Parent-offspring covariance F_2 and F_3 data
Cov(X,X'/2,1;4,4)	Analysis of variance of F_4 data
Cov(X,X'/3,2;4,4)	Analysis of variance of F_4 data
Cov(X,X'/2,1;2,4)	Parent-offspring covariance F_2 and F_4 data
Cov(X,X'/2,1;3,4)	Covariance of F_2 sub-population means in the F_3 and F_4 generations
Cov(X,X'/3,2;3,4)	Covariance within F_2 sub-populations of F_3 plant values and means of F_3 progeny

If we assume that the phenotypic value of an individual is the sum of the genotypic value of the individual plus an environmental effect, that there is no genotypic-environmental interaction, and that the environmental effects are uncorrelated, then Cov(X,X'/k,k-1;n,n'), where $n < n'$, will include only genotypic effects. Thus if

$$p = g + e ,$$

and

$$p' = g' + f ,$$

where

$$e \sim \text{NID}(0, \sigma_e^2),$$

$$f \sim NID(0, \sigma_f^2),$$

then

$$\begin{aligned} Cov(pp') &= E(pp') - E(p)E(p') \\ &= E(g+e)(g'+f) - E(g+e)\ E(g'+f) \\ &= E(gg'+gf+g'e+ef) - E(g)\ E(g') \\ &= E(gg') - E(g)E(g') \\ &= Cov(gg') \ . \end{aligned}$$

The expected mean squares of the analysis of variance of the F_3 data are shown in table 1. In this table primes are used on M_1 and M_2 to distinguish them from M_1 and M_2 of the analysis of variance of the F_4 data given in table 2.

From table 1 it can be seen that $Cov(X,X'/2,1;3,3)$ can be estimated by

$$\widehat{Cov}(X,X'/2,1;3,3) = \frac{M_1'-M_2'}{\bar{r}h} \ .$$

In table 2 the expected mean squares of the combined analysis of variance of the F_4 data are shown. It can be seen from table 2 that $Cov(X,X'/2,1;4,4)$ and $Cov(X,X'/3,2;4,4)$ can be estimated by

$$\widehat{Cov}(X,X'/2,1;4,4) = \frac{M_1-M_2}{f_3 \bar{sp}\, \bar{rh}}$$

$$\widehat{Cov}(X,X'/3,2;4,4) = \frac{M_2-M_3}{\bar{r}h} \ .$$

By the nature of the estimation of $Cov(X,X'/k,k-1;n,n')$, where $n = n'$, it is apparent under the assumption of no genotypic-environmental interaction the expected value of these variances will include only genotypic effects. Thus the expected values of the seven variances and covariances

Table 1. Expected mean squares of the analysis of variance of F_3 data[a]

Source	d.f.	Mean square	Expected mean square
Reps	$r-1$		
Entries	$f_2^{sp}-1$	M_1'	$\sigma_w^2 + \bar{h}\sigma_e^2 + r\bar{h}A$
Reps x entries	$(r-1)(f_2^{sp}-1)$	M_2'	$\sigma_w^2 + \bar{h}\sigma_e^2$
Plants in plots	$\sum_{ij}(p_{ij}-1)$		σ_w^2
Total	$\sum_{ij}p_{ij}-1$		

[a] r = number of replications = 2

f_2^{sp} = number of F_2 sub-populations in the F_3 generation = 64
the super-script sp denotes sub-population

p_{ij} = number of plants in the j^{th} plot ($j = 1,2...64$) of the i^{th} replicate ($i = 1,2$)

$\bar{h}$ = harmonic mean of plants per plot

σ_w^2 = environmental variance among plants in the same plot plus $Cov(X,X'/3,2;3,3)$

σ_e^2 = environmental variance among plots in a block

$A = Cov(X,X'/2,1;3,3)$

The degrees of freedom for reps x entries and for total were corrected for the three missing plots values which were calculated.

estimated by the F_2, F_3, and F_4 data will include only genotypic effects.

As stated previously in order to determine the presence or absence of non-allelic gene interactions, we are interested in the ratio of two adjusted covariances, where adjusted covariance was defined as 2^{k-2} Cov $(X,X'/k,k-1;n,n')$. To make meaningful comparisons between expected ratios

Table 2. Expected mean squares of the combined analysis of variance of F_4 data[a]

Source	d.f.	Mean square	Expected mean square
Sets	$s-1$		
Reps in sets	$s(r-1)$		
Entries in sets	$s(f_2^{sp}f_3^{sp}-1)$		
F_2^{sp} in sets	$s(f_2^{sp}-1)$	M_1	$\sigma_w^2 + \bar{h}\sigma_e^2 + r\bar{h}C + f_3^{sp}r\bar{h}B$
F_3^{sp} in F_2^{sp} in sets	$sf_2^{sp}(f_3^{sp}-1)$	M_2	$\sigma_w^2 + \bar{h}\sigma_e^2 + r\bar{h}C$
Reps x entries in sets	$s(r-1)(f_2^{sp}f_3^{sp}-1)$	M_3	$\sigma_w^2 + \bar{h}\sigma_e^2$
Plants in plots in sets	$\sum_{ijk}(p_{ijk}-1)$		σ_w^2
Total	$\sum_{ijk}p_{ijk}-1$		

[a] s = number of sets = 8

r = number of replications in a set = 6

f_2^{sp} = number of F_2 sub-populations in the F_4 generation in a set = 8 The super-script sp denotes sub-population

f_3^{sp} = number of F_3 sub-populations within a F_2 sub-population in the F_4 generation in a set = 4

p_{ijk} = number of plants in the k^{th} plot ($k = 1, 2 \ldots 32$) in the j^{th} replicate ($j = 1, 2 \ldots 6$) in the i^{th} set ($i = 1, 2 \ldots 8$)

$\bar{h}$ = harmonic mean of plants per plot

σ_w^2 = environmental variance among plants in the same plot plus $Cov(X,X'/4,3;4,4)$

σ_e^2 = environmental variance among plots in a block

$C = Cov(X,X'/3,2;4,4)$

$B = Cov(X,X'/2,1;4,4)$

The degrees of freedom for M_3 and total were corrected for the missing plot values which were calculated.

and estimated ratios, standard errors of the estimated ratios had to be calculated. The estimated ratios are of the form

$$\frac{\text{adjusted Cov 1}}{\text{adjusted Cov 2}} = \frac{X}{Y}$$

where Cov 1 is a particular $\widehat{Cov}(X,X'/k,k-1;n,n')$ and Cov 2 is similarly defined. An approximate variance of a ratio is given by

$$V(\frac{X}{Y}) = \frac{1}{Y^2} V(X) - \frac{2X}{Y^3} Cov(X,Y) + \frac{X^2}{Y^4} V(Y)$$

so

$$V(\frac{\text{adj. Cov 1}}{\text{adj. Cov 2}}) = \frac{1}{(\text{adj. Cov 2})^2} V(\text{adj. Cov 1})$$

$$- \frac{2 \text{ adj. Cov 1}}{(\text{adj. Cov 2})^3} Cov(\text{adj. Cov 1, adj. Cov 2})$$

$$+ \frac{(\text{adj. Cov 1})^2}{(\text{adj. Cov 2})^4} V(\text{adj. Cov 2}) \quad .$$

In order to evaluate the above variance, we need to know such terms as V(adj. Cov 1) and Cov(adj. Covl, adj. Cov 2). Now

$$V(\text{adj. Cov 1}) = V(2^{k-2} \text{ Cov 1})$$

$$= 2^{2k-4} V(\text{Cov 1})$$

and

$$Cov(\text{adj. Cov 1, adj. Cov 2}) = 2^{k-2} 2^{k'-2} Cov(\text{Cov 1, Cov 2}) .$$

Thus such terms as V(Cov 1) and Cov(Cov 1, Cov 2) had to be determined.

In $n = n'$ then Cov 1 will be a variance. The estimated variances were obtained from analysis of variance tables and were estimated by linear functions of independent mean squares. The variances of these estimated variances are linear functions of the variances of the mean squares. For example

$$V[(\widehat{Cov}(X,X'/2,1;3,3)] = V\left(\frac{M_1'-M_2'}{r\bar{h}}\right)$$

$$= \frac{1}{(r\bar{h})^2}\ [V(M_1') + V(M_2')] \ .$$

The variance of a mean square is estimated unbiasedly by

$$V(M.S.) = \frac{2(M.S.)^2}{d.f. + 2} \ .$$

The variance of $\widehat{Cov}(X,X'/2,1;4,4)$ and $\widehat{Cov}(X,X'/3,2;4,4)$ were obtained in like manner.

In the case of the ratio

$$\frac{\text{adj. } \widehat{Cov}(X,X'/2,1;4,4)}{\text{adj. } \widehat{Cov}(X,X'/3,2;4,4)}$$

two variances are estimated from the same analysis of variance table. To obtain the variance of this ratio the procedure outlined by Kempthorne (1957, p. 246) was used. In this case

$$X = M_1 - M_2$$

$$Y = 8M_2 - 8M_3$$

$$V(X) = V(M_1) + V(M_2)$$

$$V(Y) = 64V(M_2) + 64V(M_3)$$

$$Cov(X,Y) = -8V(M_2) \ .$$

When $n < n'$ Cov 1 is a covariance. The variance of an estimated covariance is given by Kendall (1952) as

$$V(Cov) = \frac{s_1^2\ s_2^2(1+r^2)}{n}$$

where

s_1^2 = estimated variance of one of the variables

s_2^2 = estimated variance of the other variable

r = estimated correlation coefficient of the two variables

n = number of observations.

By use of the above formula, variances of estimated covariances were obtained for the case where $k = 2$ and $n < n'$.

Formulas for covariances of the form Cov(Cov 1, Cov 2), where Cov 1 is a particular $\widehat{Cov}(X,X'/k,k-1;n,n')$ and Cov 2 also is so defined, were not found in the literature. Also, a formula for $V[(\widehat{Cov}(X,X'/3,2;3,4)]$ was not found. These formulas were derived in the course of this investigation and are shown in the appendix. The methods used in their derivation are also shown.

Using the equations shown above and those in the appendix $V(X)$, $V(Y)$, and $Cov(X,Y)$ were evaluated and thus $V(\frac{X}{Y})$ determined. The standard error of the ratio of two adjusted $\widehat{Cov}(X,X'/k,k-1;n,n')$ were obtained as

$$s.e. = \sqrt{V(\tfrac{X}{Y})} \ .$$

Comstock and Robinson (1948) showed that if the gene frequencies were one half then an average value of $\underline{a}$ over s loci, where $\underline{a}$, as defined previously, is a measure of the degree of dominance and can be estimated by

$$\bar{a} = \sqrt{\frac{2\sigma_D^2}{\sigma_A^2}} \ .$$

We will call $\bar{a}$ the average degree of dominance.

Horner et al. (1955) showed that non-allelic gene interaction effects will bias this estimate upward when it is obtained from a model assuming no non-allelic gene interaction effects. It was felt that it would be of interest to obtain an estimate of $\bar{a}$ based on the additive model with dominance for each of the characters measured in this study and to compare

this estimate of $\bar{a}$ to the results of the tests for non-allelic gene interaction effects.

It will be recalled that for the additive model with dominance

$$\mathrm{Cov}(X,X'/k,k-1;n,n') = \left(\frac{1}{2^{k-2}}\right)\sigma_A^2 + \left(\frac{1}{2^{n+n'-k-2}}\right)\sigma_D^2$$

so that

$$\mathrm{Cov}(X,X'/2,1;4,4) = \sigma_A^2 + \frac{1}{16}\sigma_D^2$$

and

$$\mathrm{Cov}(X,X'/3,2;4,4) = \frac{1}{2}\sigma_A^2 + \frac{1}{8}\sigma_D^2 \quad .$$

The quantities $\widehat{\mathrm{Cov}}(X,X'/3,2;4,4)$ and $\widehat{\mathrm{Cov}}(X,X'/2,1;4,4)$ were obtained from mean squares of the combined analysis of variance table of F_4 data so that $\hat{\sigma}_D^2$ and $\hat{\sigma}_A^2$ are linear functions of M_1, M_2 and M_3. Therefore

$$\frac{2\hat{\sigma}_D^2}{\hat{\sigma}_A^2} = \frac{-8M_1 + 72M_2 - 64M_3}{M_1 - 3M_2 + 2M_3} \quad .$$

The variance of the estimate of $\bar{a}$ is

$$V\sqrt{\frac{2\hat{\sigma}_D^2}{\hat{\sigma}_A^2}} \quad .$$

This is of the form $V\sqrt{Z}$. An approximate estimate of $V\sqrt{Z}$ is given by

$$V\sqrt{Z} = \frac{V(Z)}{4Z} \quad .$$

$V(Z)$ is of the form

$$V\left(\frac{X}{Y}\right) = V\left(\frac{2\hat{\sigma}_D^2}{\hat{\sigma}_A^2}\right) .$$

In this case

$$X = -8M_1 + 72M_2 - 64M_3$$

$$Y = M_1 - 3M_2 + 2M_3$$

$$V(X) = 64V(M_1) + 5184V(M_2) + 4096V(M_3)$$

$$V(Y) = V(M_1) + 9V(M_2) + 4V(M_3)$$

$$Cov(X,Y) = -8V(M_1) - 216V(M_2) - 128V(M_3) \quad .$$

The standard error for $\bar{a}$ was obtained as the square root of

$$V\sqrt{\frac{2\hat{\sigma}_D^2}{\hat{\sigma}_A^2}} \quad .$$

Kempthorne (1956b) obtained a general formula for the covariance between relatives under selfing with general epitasis. His formula is applicable to $Cov(X,X'/2,1;n,n')$ used in this thesis. Using slightly different notation than Kempthorne

$$Cov(X,X'/2,1;n,n') = \sigma_{yy} + \sum_{i=1}^{n}(K_r^i + K_s^i)(-1)^i \sigma_{yi}$$

$$+ \sum_{i,j=1}^{n} K_r^i K_s^j (-1)^{i+j} \sigma_{ij}$$

where

$$K_r = 2(1 - \frac{1}{2^r})$$

r = the number of generations of selfing to the Fn generation. In this study the F_2 is the zero generation.

$$K_s = 2(1 - \frac{1}{2^s})$$

s = the number of generations of selfing to the $F_{n'}$ generation

σ_{yy} is the genotypic variance in the F_2 generation

σ_{yi} is the covariance of the genotypic value and a value symbolized by d_i.

σ_{ii} is the variance of d_i

σ_{ij} is the covariance of d_i and d_j

d_i is a measure of the sum of the dominance deviations when $i = 1$, and a measure of epistatic effects when $i > 1$.

thus the terms σ_{yi} and σ_{ij} may be considered as terms involving epistatic effects. σ_{ii} will also involve epistatic effects when $i > 1$.

In this study five different $Cov(X,X'/2,1;n,n')$'s are estimated for each character. Therefore the data are adequate for fitting a model involving five parameters or less. By assuming all parameters in the above equation are zero except five and equating the observed covariances to expected covariances, five linear equations with five unknowns are obtained. The five parameters estimated are σ_{yy}, σ_{y1}, σ_{y2}, σ_{11}, and σ_{12}. Assuming all other parameters to be zero, the general equation reduces to

$$Cov(X,X'/2,1;n,n') = \sigma_{yy} - (K_r + K_s)\,\sigma_{y1} + (K^2_r + K^2_s)\,\sigma_{y2} + K_rK_s\,\sigma_{11} - (K_rK^2_s + K^2_rK_s)\,\sigma_{12}$$

The five linear equations are

$$Cov(X,X'/2,1;3,3) = \sigma_{yy} - 2\sigma_{y1} + 2\sigma_{y2} + \sigma_{11} - 2\sigma_{12}$$

$$Cov(X,X'/2,1;4,4) = \sigma_{yy} - 3\sigma_{y1} + 9/2\,\sigma_{y2} + 9/4\,\sigma_{11} - 27/4\,\sigma_{12}$$

$$Cov(X,X'/2,1;3,4) = \sigma_{yy} - 5/2\,\sigma_{y1} + 13/4\,\sigma_{y2} + 3/2\,\sigma_{11} - 15/4\,\sigma_{12}$$

$$Cov(X,X'/2,1;2,3) = \sigma_{yy} - \sigma_{y1} + \sigma_{y2}$$

$$Cov(X,X'/2,1;2,4) = \sigma_{yy} - 3/2\,\sigma_{y1} + 9/4\,\sigma_{y2}$$

Using these equations, estimates of σ_{yy}, σ_{y1}, σ_{y2}, σ_{11}, and σ_{12} were obtained.

The solution of the above five linear equations led to results which were unrealistic. Upon re-examination of the general formula, it was con-

cluded that a more reasonable cut off for this formula would be at three parameters instead of five. In the above five linear equations σ_{11} and σ_{12} were assumed to be zero and then estimates of σ_{yy}, σ_{y1}, and σ_{y2} obtained by the method of least squares. Next σ_{y2} and σ_{12} were assumed zero and estimates of σ_{yy}, σ_{y1}, and σ_{11} obtained, also by least squares. Comparisons were made between the estimates obtained by these two different groupings of three parameters.

EXPERIMENTAL RESULTS

In table 3 is shown estimates and their standard errors of the seven adjusted $Cov(X,X'/k,k-1;n,n')$'s which were obtained in this study for the characters kernel row number, ear length, ear diameter, and seed weight. All estimates were significantly different from zero at the five percent level except $\widehat{Cov}(X,X'/2,1;2,3)$ and $\widehat{Cov}(X,X'/2,1;2,4)$ for the character seed weight. An estimate was considered significantly different from zero if it was greater than twice its standard error. F tests in the analysis of variance tables of the F_3 and F_4 data also showed that $\widehat{Cov}(X,X'/2,1;3,3)$, $\widehat{Cov}(X,X'/2,1;4,4)$ and $\widehat{Cov}(X,X'/3,2;4,4)$ were significantly different from zero for all four characters.

Table 3. Estimated adjusted $Cov(X,X'/k,k-1;n,n')$'s and their standard errors

Covariance	Character							
	Kernel row no.		Ear length		Ear diameter		Seed weight	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
2,1;3,3	1.79*	.38	134.15*	35.21	2.48*	.61	189.39*	54.47
2,1;2,3	1.28*	.34	89.46*	33.77	3.04*	.91	77.01	53.29
2,1;4,4	1.01*	.22	134.65*	31.90	1.95*	.52	419.00*	94.48
3,2;4,4	1.10*	.16	201.05*	29.44	4.27*	.66	525.20*	71.47
2,1;2,4	1.02*	.25	106.36*	38.16	1.77*	.75	14.10	70.89
2,1;3,4	1.16*	.23	117.14*	32.07	1.75*	.44	216.38*	59.68
3,2;3,4	.90*	.09	66.00*	7.24	2.64*	.30	156.59*	16.79

*Significantly different from zero at the 5 percent level.

Estimated ratios of adjusted Cov(X,X'/k,k-1;n,n')'s and their standard errors are shown in table 4. Ratios are reported either as adj. $\widehat{Cov}$ 2/adj. $\widehat{Cov}$ 1 or adj. $\widehat{Cov}$ 1/adj. $\widehat{Cov}$ 2, whichever is less than one. This is done because the formula used to obtain the variance of a ratio does not give reliable estimates of the variance if the ratio is greater than one. For the character seed weight the ratios involving adjusted $\widehat{Cov}(X,X'/2,1;2,3)$ and adjusted $\widehat{Cov}(X,X'/2,1;2,4)$ were not determined because these two covariances were not significantly different from zero. Therefore, the ratios involving these two adjusted covariances were considered as indeterminate.

In table 5 the expected ratios of adjusted covariances for the additive model with dominance and for the four non-allelic interaction models investigated in this study are reported. Ratios are reported for 2, 8, 16, 32 and ∞ loci. It will be noted that except for one case the limit of a ratio as the number of loci goes to infinity is 0 or ∞ for the complementary, duplicate factor and multiplicative models and is 1, 0.4, or 2.5 for the optimum number model. For the ratio adjusted Cov(X,X'/2,1;2,4)/adjusted Cov(X,X'/2,1;2,3) the limit for the duplicate factor model is 1.00. The cases of partial dominance, a = 0.5, and overdominance, a = 1.5, are shown for the additive model with dominance. The expected ratio for a = 1.0, complete dominance, lies between the ratios for a = 0.5 and a = 1.5. For the multiplicative model only the case of complete dominance, e = b, is reported. The values b = 1.06 and b = 1.4 were used to provide a range which is probably sufficient for real situations. In the optimum number model d can vary from -s to +s. Because adjusted Cov(X,X'/k,k-1;n,n') for this model involves d only as d^2, values of d from zero to s except d = 32

Table 4. Estimated ratios of adjusted Cov(X,X'/k,k-1;n,n')'s and their standard errors (Ratio = adj. $\widehat{Cov}$ 2/adj. $\widehat{Cov}$ 1 or its reciprocal)

Cov 1	Cov 2	Character							
		Kernel row no.		Ear length		Ear diameter		Seed weight	
		Ratio	S.E.	Ratio	S.E.	Ratio	S.E.	Ratio	S.E.
2,1;3,3	2,1;2,3	.72	.15	.67	.24	.81[a]	.21	I[b]	
	2,1;4,4	.56[c]	.12	1.00	.28	.79	.24	.45[ad]	.14
	3,2;4,4	.62[c]	.16	.67[a]	.20	.58[ac]	.17	.36[ad]	.11
	2,1;2,4	.57[d]	.11	.79	.29	.71	.30	I	
	2,1;3,4	.65	.15	.87	.31	.71	.22	.88[a]	.32
	3,2;3,4	.51[d]	.12	.49[d]	.14	.94[a]	.26	.83	.26
2,1;2,3	2,1;4,4	.78	.18	.66[a]	.25	.64	.20	I	
	3,2;4,4	.86	.26	.44[ad]	.18	.71[a]	.24	I	
	2,1;2,4	.80	.20	.84[a]	.38	.58	.27	I	
	2,1;3,4	.90	.17	.76[a]	.27	.58	.17	I	
	3,2;3,4	.70	.20	.74	.29	.87	.28	I	
2,1;4,4	3,2;4,4	.91[a]	.25	.67[a]	.20	.46[ac]	.15	.80[a]	.22
	2,1;2,4	.98[a]	.21	.79	.22	.91	.38	I	
	2,1;3,4	.87[a]	.15	.87	.13	.90	.21	.52[d]	.12
	3,2;3,4	.90	.23	.49[d]	.13	.74[a]	.22	.37[d]	.09
3,2;4,4	2,1;2,4	.93	.26	.53[c]	.21	.41[d]	.19	I	
	2,1;3,4	.95[a]	.23	.58	.18	.41[d]	.12	.41[d]	.13
	3,2;3,4	.82[d]	.06	.33[d]	.03	.62[d]	.06	.30[d]	.02
2,1;2,4	2,1;3,4	.88[a]	.16	.91[a]	.32	.99	.37	I	
	3,2;3,4	.88	.23	.62	.23	.67[a]	.30	I	
2,1;3,4	3,2;3,4	.78	.17	.56[d]	.17	.66[a]	.18	.72	.21

[a]The reciprocal of adj. $\widehat{Cov}$ 2/adj. $\widehat{Cov}$ 1 and its standard error is given.

[b]I denotes that the ratio was indeterminate because one or both $\widehat{Cov}$ 1 and/or $\widehat{Cov}$ 2 were not significantly different from zero.

[c]The differences between the estimated ratio and the expected ratios based on an additive model and additive model with dominance are greater than 2 S.E.

[d]The differences between the estimated ratio and the expected ratios based on an additive model and additive model with dominance are greater than 2.5 S.E.

Table 5. Expected ratios of adjusted Cov(X,X'/k,k-1;n,n')'s (Ratio = adj. Cov 2/adj. Cov 1)

Cov 1	Cov 2	No. of loci	Model: Additive with dominance a=0.5	Additive with dominance a=1.5	Complementary	Duplicate factor	Multiplicative[a] b=1.06	Multiplicative[a] b=1.4	Optimum number d=0	d=2	d=8	d=16	d=∞
2,1;3,3	2,1;2,3	2	1.03	1.22	1.32	.91	1.12	1.16	1.11	1.01			
		8	1.03	1.22	3.61	.39	1.23	1.47	1.02	1.01	1.00		
		16	1.03	1.22	13.42	.15	1.30	2.02	1.01	1.00	1.00	1.00	
		32	1.03	1.22	>50.00	.02	1.46	3.81	1.00	1.00	1.00	1.00	
		∞	1.03	1.22	∞	0	∞	∞	1.00	1.00	1.00	1.00	1.00
	2,1;4,4	2	.98	.84	.76	1.11	.91	.88	.92	.99			
		8	.98	.84	.25	2.79	.89	.69	.99	1.00	1.00		
		16	.98	.84	.06	8.00	.86	.50	.99	1.00	1.00	1.00	
		32	.98	.84	<.01	>50.00	.76	.26	1.00	1.00	1.00	1.00	
		∞	.98	.84	0	∞	0	0	1.00	1.00	1.00	1.00	1.00
	3,2;4,4	2	1.00	1.00	1.04	1.69	.99	.97	2.33	1.09			
		8	1.00	1.00	1.19	22.12	.94	.81	2.48	1.45	1.04		
		16	1.00	1.00	1.13	>50.00	.90	.64	2.49	1.72	1.08	1.02	
		32	1.00	1.00	.76	>50.00	.80	.39	2.49	1.99	1.16	1.04	
		∞	1.00	1.00	0	∞	0	0	2.50	2.50	2.50	2.50	1.00
	2,1;2,4	2	1.00	1.00	1.07	.85	1.00	1.02	1.00	1.00			
		8	1.00	1.00	1.56	.39	1.03	1.14	1.00	1.00	1.00		
		16	1.00	1.00	2.49	.15	1.07	1.33	1.00	1.00	1.00	1.00	
		32	1.00	1.00	6.23	.02	1.13	1.79	1.00	1.00	1.00	1.00	
		∞	1.00	1.00	∞	0	∞	∞	1.00	1.00	1.00	1.00	1.00

[a]Only the case of complete dominance, e = b, is reported for this model.

Table 5. (Continued)

Cov 1	Cov 2	No. of loci	Model: Additive with dominance a=0.5	Additive with dominance a=1.5	Complementary	Duplicate factor	Multiplicative[a] b=1.06	Multiplicative[a] b=1.4	Optimum number d=0	Optimum number d=2	Optimum number d=8	Optimum number d=16	Optimum number d=∞
2,1;3,3	2,1;3,4	2	.98	.89	.86	1.03	.94	.93	.94	1.00			
		8	.98	.89	.48	1.53	.99	.82	.99	1.00	1.00		
		16	.98	.89	.23	2.38	.95	.69	1.00	1.00	1.00	1.00	
		32	.98	.89	.05	5.64	.90	.50	1.00	1.00	1.00	1.00	
		∞	.98	.89	0	∞	0	0	1.00	1.00	1.00	1.00	1.00
	3,2;3,4	2	1.03	1.22	1.26	1.77	1.11	1.10	2.44	1.10			
		8	1.03	1.22	2.41	16.96	1.09	1.04	2.49	1.46	1.04		
		16	1.03	1.22	4.57	>50.00	1.06	.96	2.49	1.73	1.08	1.02	
		32	1.03	1.22	12.36	>50.00	1.00	.82	2.50	1.99	1.16	1.04	
		∞	1.03	1.22	∞	∞	0	0	2.50	2.50	2.50	2.50	1.00
2,1;2,3	2,1;4,4	2	.95	.69	.58	1.23	.81	.76	.83	.99			
		8	.95	.69	.07	7.14	.73	.47	.97	.99	1.00		
		16	.95	.69	<.01	>50.00	.66	.25	.99	.99	1.00	1.00	
		32	.95	.69	<.01	>50.00	.52	.07	.99	1.00	1.00	1.00	
		∞	.95	.69	0	∞	0	0	1.00	1.00	1.00	1.00	1.00
	3,2;4,4	2	.97	.82	.79	1.86	.89	.84	2.10	1.08			
		8	.97	.82	.33	>50.00	.77	.55	2.43	1.45	1.04		
		16	.97	.82	.08	>50.00	.69	.31	2.47	1.72	1.08	1.02	
		32	.97	.82	<.01	>50.00	.55	.10	2.48	1.98	1.16	1.04	
		∞	.97	.82	0	∞	0	0	2.50	2.50	2.50	2.50	1.00

Table 5. (Continued)

Cov 1	Cov 2	No. of loci	Model: Additive with dominance a=0.5	Additive with dominance a=1.5	Complementary	Duplicate factor	Multiplicative[a] b=1.06	Multiplicative[a] b=1.4	Optimum number d=0	Optimum number d=2	Optimum number d=8	Optimum number d=16	Optimum number d=∞
2,1;2,3	2,1;2,4	2	.97	.82	.81	.94	.90	.88	.90	.99			
		8	.97	.82	.43	1.00	.84	.78	.98	1.00	1.00		
		16	.97	.82	.19	1.00	.82	.66	.99	1.00	1.00	1.00	
		32	.97	.82	.03	1.00	.77	.47	1.00	1.00	1.00	1.00	
		∞	.97	.82	0	1.00	0	0	1.00	1.00	1.00	1.00	1.00
	2,1;3,4	2	.96	.73	.65	1.14	.84	.80	.85	.99			
		8	.96	.73	.13	3.92	.80	.56	.97	.99	1.00		
		16	.96	.73	.02	15.63	.73	.34	.99	.99	1.00	1.00	
		32	.96	.73	<.01	>50.00	.61	.13	.99	1.00	1.00	1.00	
		∞	.96	.73	0	∞	0	0	1.00	1.00	1.00	1.00	1.00
	3,2;3,4	2	1.00	1.00	.96	1.95	.99	.95	2.20	1.09			
		8	1.00	1.00	.67	43.35	.89	.71	2.45	1.45	1.04		
		16	1.00	1.00	.34	>50.00	.81	.48	2.48	1.72	1.08	1.02	
		32	1.00	1.00	.07	>50.00	.69	.22	2.49	1.98	1.16	1.04	
		∞	1.00	1.00	0	∞	0	0	2.50	2.50	2.50	2.50	1.00
2,1;4,4	3,2;4,4	2	1.02	1.20	1.38	1.52	1.09	1.10	2.54	1.09			
		8	1.02	1.20	4.78	7.94	1.05	1.17	2.51	1.46	1.04		
		16	1.02	1.20	18.96	45.88	1.05	1.28	2.50	1.73	1.08	1.02	
		32	1.02	1.20	>50.00	>50.00	1.05	1.50	2.50	1.99	1.16	1.06	
		∞	1.02	1.20	∞	∞	∞	∞	2.50	2.50	2.50	2.50	1.00

Table 5. (Continued)

Cov 1	Cov 2	No. of loci	Model: Additive with dominance a=0.5	Additive with dominance a=1.5	Complementary	Duplicate factor	Multiplicative[a] b=1.06	Multiplicative[a] b=1.4	Optimum number d=0	Optimum number d=2	Optimum number d=8	Optimum number d=16	Optimum number d=∞
2,1;4,4	2,1;2,4	2	1.02	1.20	1.41	.77	1.10	1.16	1.09	1.01			
		8	1.02	1.20	6.25	.14	1.15	1.65	1.01	1.00	1.00		
		16	1.02	1.20	41.75	.02	1.24	2.67	1.01	1.00	1.00	1.00	
		32	1.02	1.20	>50.00	<.01	1.48	6.92	1.00	1.00	1.00	1.00	
		∞	1.02	1.20	∞	0	∞	∞	1.00	1.00	1.00	1.00	1.00
	2,1;3,4	2	1.01	1.06	1.13	.93	1.03	1.05	1.03	1.00			
		8	1.01	1.06	1.94	.55	1.10	1.19	1.00	1.00	1.00		
		16	1.01	1.06	3.83	.30	1.11	1.39	1.00	1.00	1.00	1.00	
		32	1.01	1.06	14.50	.09	1.17	1.93	1.00	1.00	1.00	1.00	
		∞	1.01	1.06	∞	0	∞	∞	1.00	1.00	1.00	1.00	1.00
	3,2;3,4	2	1.05	1.46	1.67	1.59	1.22	1.25	2.67	1.10			
		8	1.05	1.46	9.71	6.06	1.22	1.51	2.53	1.47	1.04		
		16	1.05	1.46	>50.00	20.32	1.23	1.93	2.51	1.73	1.08	1.02	
		32	1.05	1.46	>50.00	>50.00	1.32	3.17	2.51	1.99	1.16	1.04	
		∞	1.05	1.46	∞	∞	∞	∞	2.50	2.50	2.50	2.50	1.00
3,2;4,4	2,1;2,4	2	1.00	1.00	1.03	.51	1.01	1.05	.43	.92			
		8	1.00	1.00	1.31	.02	1.10	1.41	.40	.69	.96		
		16	1.00	1.00	2.20	<.01	1.19	2.09	.40	.58	.92	.98	
		32	1.00	1.00	8.17	<.01	1.40	4.61	.40	.50	.86	.96	
		∞	1.00	1.00	∞	0	∞	∞	.40	.40	.40	.40	1.00

Table 5. (Continued)

Cov 1	Cov 2	No. of loci	Model: Additive with dominance a=0.5	Additive with dominance a=1.5	Complementary	Duplicate factor	Multiplicative[a] b=1.06	Multiplicative[a] b=1.4	Optimum number d=0	Optimum number d=2	Optimum number d=8	Optimum number d=16	Optimum number d=∞
3,2;4,4	2,1;3,4	2	.98	.89	.82	.61	.95	.95	.40	.92			
		8	.98	.89	.40	.07	1.05	1.01	.40	.69	.96		
		16	.98	.89	.20	.01	1.06	1.09	.40	.58	.92	.98	
		32	.98	.89	.07	<.01	1.11	1.28	.40	.50	.86	.96	
		∞	.98	.89	0	0	∞	∞	.40	.40	.40	.40	1.00
	3,2;3,4	2	1.03	1.22	1.21	1.05	1.12	1.13	1.05	1.01			
		8	1.03	1.22	2.02	.77	1.16	1.29	1.01	1.00	1.00		
		16	1.03	1.22	4.05	.53	1.18	1.51	1.00	1.00	1.00	1.00	
		32	1.03	1.22	16.21	.27	1.25	2.11	1.00	1.00	1.00	1.00	
		∞	1.03	1.22	∞	0	∞	∞	1.00	1.00	1.00	1.00	1.00
2,1;2,4	2,1;3,4	2	.98	.89	.80	1.21	.94	.91	.94	1.00			
		8	.98	.89	.31	3.92	.96	.72	.99	1.00	1.00		
		16	.98	.89	.09	15.63	.89	.52	1.00	1.00	1.00	1.00	
		32	.98	.89	.01	>50.00	.79	.28	1.00	1.00	1.00	1.00	
		∞	.98	.89	0	∞	0	0	1.00	1.00	1.00	1.00	1.00
	3,2;3,4	2	1.03	1.22	1.18	2.07	1.10	1.08	2.44	1.10			
		8	1.03	1.22	1.55	43.39	1.06	.91	2.49	1.46	1.04		
		16	1.03	1.22	1.83	>50.00	.99	.72	2.49	1.73	1.08	1.02	
		32	1.03	1.22	1.98	>50.00	.89	.46	2.50	1.99	1.16	1.04	
		∞	1.03	1.22	2.00	∞	0	0	2.50	2.50	2.50	2.50	1.00
2,1;3,4	3,2;3,4	2	1.04	1.37	1.47	1.72	1.18	1.19	2.59	1.10			
		8	1.04	1.37	5.00	11.06	1.11	1.27	2.51	1.46	1.04		
		16	1.04	1.37	19.99	>50.00	1.11	1.39	2.51	1.73	1.08	1.02	
		32	1.04	1.37	>50.00	>50.00	1.12	1.64	2.50	1.99	1.16	1.04	
		∞	1.04	1.37	∞	∞	∞	∞	2.50	2.50	2.50	2.50	1.00

for s = 32 are reported in table 5. For this model the expected ratio of d = 32 at s = 32 is essentially the same as that for d = 16 at s = 32.

Comparisons were made between the expected ratios in table 5 for the additive model with dominance and the estimated ratios given in table 4. Comparisons also were made between the estimated ratios and the value 1.00 which is the expected ratio for the additive model. If the differences between the estimated ratio and the expected ratios based on the additive model and the additive model with dominance were greater than two or two and one-half times the standard error of the estimated ratio, the differences were considered as significant or highly significant respectively and so indicated in table 4. These significant differences constitute evidence for the presence of non-allelic gene interaction effects. For all four characters significant differences were found. Out of the 21 estimated ratios for kernel row number, ear length, and ear diameter, five, six, and five, respectively, significant or highly significant differences were found. Six such differences were found among the 10 estimated ratios for the character seed weight.

Having evidence from the data in table 4 that non-allelic gene interaction was operative in controlling the characters kernel row number, ear length, ear diameter, and seed weight in the cross of two inbred lines of corn, Hy and B10, an attempt was made to determine the nature of this gene action and the number of loci involved. If a particular type of epistatic gene action involving n loci was operative in controlling a character, then the estimated ratio of two adjusted covariances for that character should not be significantly different from the expected ratio for n loci, except in cases of chance occurrence of such a difference.

In table 6 is recorded the frequency of fit of the estimated ratios that were significantly different from an additive model and additive models with dominance to the four epistatic models for various numbers of loci. An estimated ratio was considered as fitting an epistatic model if the expected ratio fell within the estimated ratio's 95 percent confidence interval.

From the data in table 6 no clear cut evidence is presented to indicate that the gene action for any of the characters was one of the four epistatic types studied in this thesis. The evidence is more clear as to which type probably was not operative. For the character kernel row number there is no evidence for gene action of the optimum number type, no evidence for complementary gene action for ear length, and no evidence for multiplicative gene action for the case of complete dominance, where e = b, for seed weight. For the character ear length two out of the six estimated ratios did not fit any of the four epistatic models and one out of six did not fit any of these models for character seed weight.

It will be noted from table 5 that for a particular ratio of adjusted covariances the expected ratios for different models may take values over the same range. Therefore it is possible for an estimated ratio to fit several genetic models even though only one model is applicable. On the basis of this reasoning, one would expect that if the gene action was of only one type that the fit to a model of this type would occur in all or nearly all comparisons made between estimated and expected ratios and would fit other models only in the cases where their expected ratios happen to be in the same range as the expected ratio of the model that is operative. The data of table 7 are presented in order to look at the results of this

Table 6. Frequency of fit of estimated ratios to epistatic models[a]

Character	No. of significant ratios	No. of loci	Model								Fitted none of the four models
			Complementary	Duplicate factor	Multiplicative		Optimum number				
					b=1.06	b=1.4	d=0	d=2	d=8	d=16	
Kernel Row No.	5	2-∞	2	2	3	3					
		2-8	1	2		2					
		8-16		2	1	2					
		16-32	1		2	2					
		32-∞	1		3	2					
Ear Length	6	2-∞		3	1	1					2
		2-8		2			2	2			
		8-16				1	2	2	1		
		16-32		1		1	2	2	1		
		32-∞			1	1	2	2	2	2	
Ear Diameter	5	2-∞	3	5	1	1					
		2-8	3	4			4	2			
		8-16	3	1			4	4			
		16-32	2	1		1	4	4	1		
		32-∞			1	1	4	4	4	4	
Seed Weight	6	2-∞	1	5							1
		2-8	1	4			2				
		8-16	1	2			2	2			
		16-32	1	2			2	2			
		32-∞		1			2	2	2	2	

[a]Only the estimated ratios which were significantly different from an additive model and additive model with dominance are considered in this table. An estimated ratio was considered fitted to an epistatic model if the expected ratio fell within the estimated ratio's 95 percent confidence interval.

Table 7. Epistatic models that were fitted by estimated ratios[a]

Ratios[b]		Character			
Cov 1	Cov 2	Kernel row no.	Ear length	Ear diameter	Seed weight
2,1;3,3	2,1;4,4	C,M			DF
	3,2;4,4	C,M		C,DF,ON	DF,ON
	2,1;2,4	DF			
	3,2;3,4	M	M		
2,1;2,3	3,2;4,4		DF,ON		
2,1;4,4	3,2;4,4			C,DF,M,ON	
	2,1;3,4				DF
	3,2;3,4		None of the four		None of the four
3,2;4,4	2,1;2,4		DF,ON	DF,ON	
	2,1;3,4			C,DF,ON	C,DF,ON
	3,2;3,4	DF	DF	DF	DF
2,1;3,4	3,2;3,4		None of the four		

[a]Only estimated ratios that were significantly different from the expected ratios of the additive and additive with dominance models are considered in this table. A blank space indicates that such differences for that ratio were not significantly different for that particular character. An estimated ratio was considered as fitting an epistatic model if the expected ratio fell within the estimated ratio's 95% confidence interval.

C = complementary model DF = duplicate factor model

M = multiplicative model ON = optimum number model

[b]Ratio = adj. $\hat{Cov}$ 1/adj. $\hat{Cov}$ 2 or adj. $\hat{Cov}$ 2/adj. $\hat{Cov}$ 1 whichever was less than one. See table 4.

study in this manner. Here are shown the epistatic models that are fitted by an estimated ratio. Only estimated ratios that were significantly different from the expected ratios of the additive and additive with dominance models are considered in the table.

The estimated ratios for the character kernel row number did not consistently fit any one of the four epistatic models. There appears to be no consistency of fit to any one model for ear length, although the duplicate factor model appears three times out of five. For this character, two of the estimated ratios did not fit any of the four models. This would lend support to the hypothesis that another non-allelic gene interaction model other than the four investigated here was operative in the inheritance of this character.

The data in table 7 for the character ear diameter gives some evidence that the controlling gene action was of a duplicate factor nature. This model was fitted by all five of the estimated ratios, whereas the optimum number, complementary, and multiplicative models were fitted by four, three, and one respectively of the estimated ratios. Except for one case, whenever the values for the expected ratio of the duplicate factor model was not in the same range as that of another model, then the estimated ratio fit the duplicate factor model but not the other model.

The estimated ratios for seed weight consistently fit the duplicate factor model. This model was fitted by five out of the six estimated ratios. The one estimated ratio that did not fit the duplicate factor model also did not fit any of the other three models. None of the other models were consistently fitted for this character. The results for this character suggest that the duplicate factor model may have been operative

in controlling this character.

In table 8 are shown the estimates of $\bar{a}$, the average degree of dominance, which were estimated from the F_4 data. Also shown are the standard errors of the estimates. If the estimate of $\bar{a}$ was greater than twice its standard error, it was considered that the estimate was significantly different from zero. Although all estimates were greater than one and thus in the overdominance range, only the estimate for ear diameter was significantly different from zero. None were significantly different from one; therefore there was no evidence for overdominance for any of the four characters.

Table 8. Estimates of $\bar{a}$ and their standard errors

Character	$\bar{a}$	s.e.
Kernel row no.	1.03	1.64
Ear length	2.51	1.33
Ear diameter	4.59*	2.29
Seed weight	1.72	1.27

*Significantly different from zero at the 5 percent probability level.

Estimates for each of the four characters of the five parameters, σ_{yy}, σ_{y1}, σ_{y2}, σ_{11}, and σ_{12}, that were described by Kempthorne (1956b), are shown in table 9. σ_{yy} is the genotypic variance in the F_2 generation, σ_{11} is a measure of the dominance variance, and σ_{y1}, σ_{y2}, and σ_{12} are covariances involving epistatic effects. All estimates were negative except σ_{yy} for ear diameter. These results are rather unrealistic because

Table 9. Estimates of σ_{yy}, σ_{y1}, σ_{y2}, σ_{11}, and σ_{12}

Parameter	Character			
	Kernel row no.	Ear length	Ear diameter	Seed weight
σ_{yy}	-2.60	-229.00	2.12	-378.37
σ_{y1}	-6.82	-508.23	-3.23	-842.84
σ_{y2}	-2.94	-189.77	-2.31	-387.46
σ_{11}	-6.90	-548.34	-4.93	-1,040.03
σ_{12}	-1.76	-137.29	-1.72	-348.52

σ_{yy} and σ_{11} are variances and can only be zero or greater, but of the eight estimates of these terms seven were negative. Evidence from the F_3 and F_4 data indicates that σ_{yy} was not zero because highly significant genetic variability was detected in both the F_3 and F_4 generations for all four characters. In general the estimates of σ_{11} were of such large negative values to cast doubts as to their being estimates of zero.

After a re-examination of Kempthorne's (1956b) general formula for $Cov(X,X'/2,1;n,n')$, it was concluded that a more reasonable cut off for this formula would be at three parameters instead of five. All parameters were assumed to be zero except σ_{yy}, σ_{y1}, and σ_{y2} and then estimates obtained of these parameters. These estimates are shown in table 10. The estimates of σ_{yy}, the F_2 genotypic variance, were all positive except for seed weight. Thus it appears that the estimates in this table are more realistic than those reported in table 9.

In table 11 are shown estimates of σ_{yy}, σ_{y1}, and σ_{11}. These were

Table 10. Estimates of σ_{yy}, σ_{y1}, and σ_{y2}[a]

Parameter	Character			
	Kernel row no.	Ear length	Ear diameter	Seed weight
σ_{yy}	.99	67.17	2.75	-183.81
σ_{y1}	-1.07	-43.48	-.43	-210.54
σ_{y2}	-.72	-14.53	-.55	-20.81
$\hat{\sigma}^2_e$	.02	154.80	.28	8887.02

[a]All σ_{yi} terms where i > 2, and all σ_{ij} terms were assumed to be zero.

Table 11. Estimates of σ_{yy}, σ_{y1}, and σ_{11}[a]

Parameter	Character			
	Kernel row no.	Ear length	Ear diameter	Seed weight
σ_{yy}	2.12	71.92	5.78	226.46
σ_{y1}	.66	-23.51	2.68	151.95
σ_{11}	.48	-2.74	1.86	274.34
$\hat{\sigma}^2_e$	.18	225.30	.02	1507.12

[a]All σ_{yi} terms where i > 1, and all σ_{ij} terms where i or j > 1 were assumed to be zero.

obtained by assuming all other parameters in the general formula for $Cov(X,X'/2,1;n,n')$ to be zero. These estimates also appear to be more realistic than those given in table 9. The estimates of the two variances, σ_{yy} and σ_{ll}, were all positive except $\hat{\sigma}_{ll}$ for ear length and this was only a small negative value.

In tables 10 and 11 $\hat{\sigma}_e^2$ is a measure of the fit of estimated parameters to the experimental data. In table 10 the estimates of σ_{yy}, σ_{yl}, and σ_{y2} were substituted back into the estimating formulas and estimates of the $Cov(X,X'/2,1;n,n')$'s obtained. These estimates of the $Cov(X,X'/2,1;n,n')$'s were then compared to those obtained from the experimental data and differences calculated. $\hat{\sigma}_e^2$ is the estimated variance of these differences. Large values of $\hat{\sigma}_e^2$ indicate that parameters other than those estimated probably should be included to explain the data, whereas small values of $\hat{\sigma}_e^2$ indicate that most of genetic variance can be explained by the three parameters. No measure is given to determine if the magnitude of $\hat{\sigma}_e^2$ is large or small. Comparisons can be made between the values of $\hat{\sigma}_e^2$ in the two tables.

DISCUSSION

In estimating the true values of the adjusted $Cov(X,X'/k,k-1;n,n')$'s the following assumptions were made: (1) normal diploid behavior at meiosis, (2) no multiple alleles, (3) no linkage, (4) no selection, (5) symmetry, and (6) no genotype-environmental interaction. The estimates are unbiased estimates of the true values only in so far as these assumptions are valid.

Assumptions (1) and (2) probably were not important sources of errors in the estimates. The material used in this study consisted of selfed generations derived from a cross of two inbred lines of corn, Hy and B10. Both lines are diploids and no abnormal meiotic behavior has been noted in them. Multiple alleles could have arisen only if some of the loci were not homozygous in the two lines or by mutation in the F_1, F_2 or F_3 generation. The lines had been selfed for eight or more generations and thus should have been essentially homozygous at all loci. From what is known about mutation rates in corn, very few multiple alleles probably arose from this source.

The assumption of no linkage could have been a serious source of bias in the estimates obtained. The effect of linkage on the models used in this study has not been worked out. Linkage would have affected epistatic variance only if the interacting genes were linked. To conceive of such linkages not existing for the characters studied here would be hard to accept because of the large number of loci probably involved in the inheritance of these characters.

In so far as possible no intentional selection was practiced in this

material. How much unintentional selection was practiced could not be ascertained, but one probably is safe in assuming this was not an important source of error. Doubts as to the size of samples being large enough to reduce the sampling errors to manageable proportions may be raised. Limits of space, funds, and time are the determining factors as to size of sample in many such studies, and this often limits the sample to a size smaller than the desired one for genetic purposes.

Symmetry was defined by Horner et al. (1955) as the condition where "all genotypes having the same number of (++), (+-), and (--) loci, respectively, have equal value or effect." Whether such conditions exist in actual situations would be hard to determine.

The assumption of no genotype-environmental interaction is very likely to be unrealistic and could have caused a serious bias in the results. Sprague and Federer (1951), Byrd (1955), and others have shown genotype-environmental interaction to be of importance in corn. Gamble (1957) found that nonepistatic effects were relatively stable over environments, whereas the epistatic effects interacted with environment.

Under the assumption of unbiased estimates, the results of this study indicate that in population investigated here non-allelic gene interaction effects were present for all characters studied. It was assumed that the estimated adjusted covariance ratios were unbiased estimates of the true ratio and were distributed normally about the true estimates. It was further assumed that the standard errors of the estimated ratios were unbiased estimates of the standard deviations of the above distributions. Under these assumptions, one would expect to get a significant difference between the expected and estimated ratios five percent of the time and a highly

significant difference one percent of the time when no difference existed. However, the data of table 4 show 24, 29, 24, and 60 percent significant differences between expected and estimated ratios were obtained for kernel row number, ear length, ear diameter, and seed weight, respectively.

Evidence for the presence of non-allelic gene interaction effects in other populations of corn have been reported by Stringfield (1950), Sentz et al. (1954), Jinks (1955), Pollak et al. (1957), Gamble (1957) and others. The results of this study along with the evidence reported by other workers indicate that epistatic effects are commonly found in quantitative characters of corn and need to be taken into account when planning breeding programs.

In this study the nature of the non-allelic gene interaction involved in the inheritance of the characters studied was investigated. From the data of table 6 and 7, there is little evidence that the gene action controlling the inheritance of kernel row number or ear length was any one of the four epistatic types investigated in this study. If only one of these four types was operative in these two characters, the data may have been insufficient to detect it. Even with more data it may not have been possible to detect it if sampling errors were too large. It may have been possible that another non-allelic gene interaction model, other than the four considered here, was operative in the inheritance of these two characters. For the character ear length there was some evidence for this hypothesis in that two out of five estimated ratios fit none of the four epistatic models which were considered in this thesis.

Another possible explanation for the failure to find evidence for any one of the four types for kernel row number and ear length is that the gene

action may have been of several types. Such a possibility for the inheritance of quantitative characters was suggested by Cockerham (1959).

There was some evidence that the gene action controlling ear diameter and seed weight may have been of the duplicate factor type. The evidence for seed weight appeared to be more substantial than that for ear diameter. The evidence does not rule out the possibility that the gene action may have been of a type other than those investigated in this study or that several types were operative. Notwithstanding these possibilities, the data for ear diameter and seed weight fit the hypothesis of duplicate factor gene action reasonably well so that other hypotheses are not needed to explain the data.

If the duplicate factor model was operative for ear diameter and seed weight, then from the data of table 6 it could be concluded that the number of interacting loci involved in the inheritance of ear diameter probably was 2-8 and not over 32 and that for seed weight the number was 2-32. The literature contains little data giving estimates on the number of genes controlling quantitative characters, but the above estimates appear to be lower than that commonly accepted. Srb and Owen (1957), p. 320, state that, "geneticists have been led to believe that many quantitative characters represent the composite influence of genes at more than 10--perhaps occasionally more than 200--loci."

The upper limit of b, i.e. $b = 1.4$, for the multiplicative model was determined by Horner et al. (1955) on the assumption that the genotypic coefficient of variation would not exceed 30 percent in real situations. After examining the models of Horner et al, Cockerham (1959) concluded that it would be impossible to distinguish between an additive and a multiplica-

tive model by the methods described by Horner et al. and used here if the genotypic coefficient of variation was limited to less than 40 percent. This may be the reason that there was little evidence for the presence of multiplicative gene action in this study.

An understanding of the nature of the gene action in the inheritance of a character is important in the planning of a breeding program that involves that character. Two breeding methods that take into account epistatic effects are recurrent selection for specific combining ability and reciprocal recurrent selection. One of these methods might well be the best method to use in a breeding program with this material providing epistatic effects are present.

The method used in this study to detect the presence of non-allelic gene interaction effects apparently was effective in detecting such effects in the population considered providing the underlying assumptions do not invalidate the applicability of the argumentation. The assumptions that appear to be the most likely not to have been met are the absence of linkage of interacting loci, the absence of genotype-environmental interaction, and nonsymmetry. Before this method of detecting epistatic effects can be more intelligently evaluated, the effects of linkage on the models employed must be ascertained, techniques developed to determine the bias due to genotype-environmental interaction effects in the estimates of $Cov(X,X'/k,k-1;n,n')$, and how valid is the assumption of symmetry in real situations must be determined. Perhaps nonsymmetrical models would fit the data without epistasis.

Horner et al. (1955) showed that epistatic effects would bias upward any estimate of $\bar{a}$, average degree of dominance, if $\bar{a}$ was estimated by

models assuming no epistatic effects. Based on nonepistatic models, estimates of $\bar{a}$ were obtained in this study. Although all estimates were greater than one and thus lying in the overdominance range, statistical tests indicated that these estimates were not significantly different from zero, no dominance, in the case of kernel row number, ear length, and seed weight, and in the case of ear diameter was not significantly different from 0.5, partial dominance. Therefore, no attempt was made to compare the estimates of $\bar{a}$ with the evidence for epistatic effects.

Kempthorne (1956b) gave a general formula for the covariance between relatives under selfing with general epistasis. This formula was applicable to the Cov $(X,X'/2,1;n,n')$'s estimated in this study. The data was adequate to estimate only five or less parameters from Kempthorne's general formula. The parameters estimated were σ_{yy}, σ_{y1}, σ_{y2}, σ_{11}, and σ_{12}. σ_{y1}, σ_{y2}, and σ_{12} are covariances and involve non-allelic gene interaction effects. σ_{yy} is the genotypic variance of the F_2 population and σ_{11} is a variance which is a measure of dominance deviations.

Estimating all five parameters by solving five linear equations led to results which were unrealistic. It was decided that a more reasonable cut off for the general formula would be at three parameters instead of five. Two different groupings of three parameters were made and estimates of the parameters made. These estimates are given in tables 10 and 11. In one group σ_{yy}, σ_{y1}, and σ_{y2} were estimated and in the other group σ_{yy}, σ_{y1}, and σ_{11} were estimated. These estimates gave results which were more realistic than those obtained with five parameters.

The $\hat{\sigma}^2_e$'s reported in tables 10 and 11 are measures of the fit of the estimates of the parameters to the experimental data. A small value of $\hat{\sigma}^2_e$

indicates a good fit whereas a large value indicates a poor fit. A comparison of the $\hat{\sigma}_e^2$ in the two tables shows that $\hat{\sigma}_e^2$ is smaller in table 10 than in table 11 for kernel row number and ear length but is larger for ear diameter and seed weight. This would indicate that a model including σ_{yy}, σ_{y1}, and σ_{y2} would fit the data better for kernel row number and ear length than one including σ_{yy}, σ_{y1}, and σ_{11}, the opposite would be true for ear diameter and seed weight. This would indicate that dominance effects were relatively more important in ear diameter and seed weight than in kernel row number and ear length.

SUMMARY

1. Investigations were conducted as to the nature of the gene action in the inheritance of four quantitative characters in a population of corn derived from a single cross of two homozygous inbred lines. The characters studied were number of kernel rows, ear length, ear diameter, and total seed weight. The material was selfed to the F_4 generation. Single ears from 64 random plants were harvested in the F_2 generation. Progenies of these 64 plants were grown in a replicated trial in the F_3 generation. Progenies of four F_3 plants within each of the 64 F_2 populations were grown in a replicated trial in the F_4 generation.

2. A method for detecting non-allelic gene interaction effects in populations derived by selfing was described. The method consists of making comparisons of estimated genetic ratios with those expected based on nonepistatic models. Upon finding evidence for epistatic effects the nature of the epistatic gene action may be investigated by making comparisons of the estimated ratios with expected ratios based on non-allelic gene interaction models. Four non-allelic gene interaction models were considered. These were the complementary, duplicate factor, multiplicative, and optimum number models.

3. From the F_2, F_3, and F_4 data, seven genetic variances and covariances were estimated for each character. From these, various genetic ratios were estimated. The estimated ratios were compared to those expected based on nonepistatic models. Evidence was found for the presence of non-allelic gene interaction effects in all characters. The data for ear diameter and seed weight fit the duplicate factor model indicating this type of gene

action may have been operative in the inheritance of these two characters. No consistent fit of the data to any of the four epistatic models was found for kernel row number and seed weight. Estimates were made of the number of loci involved in the inheritance of ear diameter and seed weight.

4. The conclusions reached in this study were based on certain assumptions. The assumptions most likely not to have been met were symmetry, no linkage, and no genotypic-environmental interaction. Nonvalidity of these assumptions could have caused serious bias in the results giving evidence for the presence of non-allelic gene interaction effects when such effects were not present.

5. Estimates of the average degree of dominance were obtained for all four characters. These estimates were obtained from a nonepistatic model. No evidence was found for overdominance.

6. Estimates of parameters designated as σ_{yy}, σ_{y1}, σ_{y2}, σ_{11}, and σ_{12} were obtained. σ_{yy} is the genotypic variance of the F_2 generation, σ_{11} is a variance which is a measure of the dominance deviations, and σ_{y1}, σ_{y2}, and σ_{12} are covariances involving epistatic effects. Unrealistic estimates were obtained when all five parameters were estimated. More realistic results were obtained by fitting the data to only three parameters. Two different groupings of three parameters were made and estimates obtained of the parameters. Comparisons between the two different groupings were made.

7. In order to evaluate the standard errors of the estimated genetic ratios expectation of covariances of the form Cov(Cov 1, Cov 2) had to be derived. Cov 1 and Cov 2 are estimated genetic variances or covariances.

The method used in the derivation of these expectations was shown and the results reported.

LITERATURE CITED

Allard, R. W. 1956a. Estimation of prepotency from lima bean diallel cross data. Agron. Jour. 48: 537-543.

_____. 1956b. The analysis of genetic-environmental interactions by means of diallel crosses. Genetics 41: 305-318.

Anderson, V. L. and Kempthorne, O. 1954. A model for the study of quantitative inheritance. Genetics 39: 883-898.

Bateson, W. 1907. Facts limiting the theory of heredity. Science 26: 649-660.

_____ and Punnett, R. C. 1906. Comb characters. Reports to Evolution Com. Roy. Soc. 3: 11-16.

_____, Saunders, E. R., and Punnett, R. C. 1905. Sweet pea (Lathyrus odoratus). Reports to Evolution Com. Roy. Soc. 2: 80-99.

_____, _____, and _____. 1906. Flower-colour in sweet peas (Lathyrus) and stocks (Matthiola). Reports to Evolution Com. Roy. Soc. 3: 3-5.

Bauman, L. F. 1959. Evidence of non-allelic gene interaction in determining yield, ear height, and kernel row number in corn. Agron. Jour. 51: 531-534.

Byrd, W. P. 1955. Genetic and environmental variances in segregating and non-segregating maize populations. Unpublished Ph.D. Thesis. Ames, Iowa, Library, Iowa State University of Science and Technology.

Charles, D. R. and Smith, H. H. 1939. Distinguishing between two types of gene action in quantitative inheritance. Genetics 24: 34-48.

Cockerham, C. C. 1954. An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. Genetics 39: 859-882.

_____. 1956. Effects of linkage on the covariances between relatives. Genetics 41: 138-141.

_____. 1959. Partitions of hereditary variance for various genetic models. Genetics 44: 1141-1148.

Comstock, R. E. and Robinson, H. F. 1948. The components of genetic variance in populations of biparental progenies and their use in estimating the average degree of dominance. Biometrics 4: 254-266.

_____ and _____. 1952. Estimation of average dominance of genes. In Gowen, J. W., ed. Heterosis, pp. 494-516. Ames, Iowa, Iowa State College Press.

Dickinson, A. G. and Jinks, J. L. 1956. A generalized analysis of diallel crosses. Genetics 41: 65-78.

East, E. M. 1910. A Mendelian interpretation of variation that is apparently continuous. Amer. Nat. 44: 65-82.

_____. 1913. Inheritance of flower size in crosses between species of Nicotiana. Bot. Gaz. 55: 177-188.

Fisher, R. A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. Trans. Roy. Soc. of Edinburgh 52: 399-433.

_____, Immer, F. R., and Tedin, O. 1932. The genetical interpretation of statistics of the third degree in the study of quantitative inheritance. Genetics 17: 107-124.

Gamble, E. E. 1957. Relative importance of epistatic and nonepistatic gene effects in maize. Unpublished Ph.D. Thesis. Ames, Iowa, Library, Iowa State University of Science and Technology.

Grafius, J. E. 1959. Heterosis in barley. Agron. Jour. 51: 551-554.

Griffing, B. 1950. Analysis of quantitative gene action by constant parent regression and related techniques. Genetics 35: 303-321.

_____. 1956. A generalized treatment of the use of diallel crosses in quantitative inheritance. Heredity 10: 31-50.

Hayman, B. I. 1954a. The analysis of variance of diallel tables. Biometrics 10: 235-244.

_____. 1954b. The theory and analysis of diallel crosses. Genetics 39: 789-809.

_____. 1957. Interaction, heterosis, and diallel crosses. Genetics 42: 336-355.

_____. 1958a. The theory and analysis of diallel crosses. II. Genetics 43: 63-85.

_____. 1958b. The separation of epistatic from additive and dominance variation in generation means. Heredity 12: 371-390.

_____ and Mather, K. 1955. The description of genic interactions in continuous variation. Biometrics 11: 69-82.

Hollander, W. F. 1955. Epistasis and hypostasis. Jour. Heredity 46: 222-225.

Horner, T. W. 1956a. Full sib and parent offspring correlations in symmetrical random mating populations with the frequency of the more favorable allele the same at all loci. Iowa State College Jour. Sci. 31: 55-65.

_____. 1956b. The components of variance in symmetrical random mating populations with the frequency of the more favorable allele the same at all loci. Iowa State Jour. Sci. 31: 67-77.

_____, Comstock, R. E., and Robinson, H. F. 1955. Non-allelic gene interactions and the interpretation of quantitative genetic data. North Carolina Agr. Exp. Sta. Tech. Bul. 118.

_____ and Kempthorne, O. 1955. The components of variance and the correlations between relatives in symmetrical random mating populations. Genetics 40: 310-320.

_____ and Weber, C. R. 1956. Theoretical and experimental study of self fertilized populations. Biometrics 12: 404-414.

Jinks, J. L. 1954. The analysis of continuous variation in a diallel cross of Nicotiana rustica varieties. Genetics 39: 767-788.

_____. 1955. A survey of the genetical basis of heterosis in a variety of diallel crosses. Heredity 9: 223-238.

_____. 1956. The F_2 and backcross generations from a set of diallel crosses. Heredity 10: 1-30.

_____ and Morley Jones, R. 1958. Estimation of the componants of heterosis. Genetics 43: 223-234.

_____ and Stevens, J. M. 1959. The components of variation among family means in diallel crosses. Genetics 44: 297-308.

Kempthorne, O. 1954. The correlations between relatives in a random mating population. Proc. Roy. Soc. London, B. 143: 103-113.

_____. 1955. The theoretical values of correlations between relatives in random mating populations. Genetics 40: 153-167.

_____. 1956a. The theory of the diallel cross. Genetics 41: 451-459.

_____. 1956b. On the covariances between relatives under selfing with general epistacy. Proc. Roy. Soc. London, B. 149: 100-108.

_____. 1957. An introduction to genetic statistics. New York, N. Y., John Wiley and Sons.

Kendall, M. G. 1952. The advanced theory of statistics. Vol. 1, 5th ed. London, Griffin.

Khambanonda, I. 1948. Polygenic inheritance of fruit size in red pepper. Unpublished Ph.D. Thesis. Ames, Iowa, Library, Iowa State University of Science and Technology.

_____. 1950. Quantitative inheritance of fruit size in red pepper. (Capsicum frutescens L.) Genetics 35: 322-343.

Kinman, M. L. and Sprague, G. F. 1945. Relation between number of parental lines and theoretical performance of synthetic varieties of corn. Jour. Amer. Soc. Agron. 37: 341-351.

Lindstrom, E. W. 1935. Segregation of quantitative genes for tetraploid tomato hybrids as evidence for dominance relations in size characters. Genetics 20: 1-11.

MacArthur, J. W. and Butler, L. 1938. Size inheritance and geometric growth processes in the tomato fruit. Genetics 23: 253-268.

Mather, K. 1949. Biometrical genetics. New York, N. Y., Dover Publications, Inc.

_____ and Vines, A. 1952. The inheritance of height and flowering time in a cross of Nicotiana rustica. In Reeve, E. C. R., and Waddington, C. H., eds. Quantitative inheritance. pp. 49-79. London, Her Majesty's Stationery Office.

Matzinger, D. F. 1956. Components of variance of diallel crosses of maize in experiments repeated over locations and years. Unpublished Ph.D. Thesis. Ames, Iowa, Library, Iowa State University of Science and Technology.

_____ and Kempthorne, O. 1956. The modified diallel table with partial inbreeding and interactions with environment. Genetics 41: 822-833.

Neal, N. P. 1935. The decrease in yielding capacity in advanced generations of hybrid corn. Jour. Amer. Soc. Agron. 27: 666-670.

Nilsson-Ehle, H. 1909. Kreuzungsuntersuchungen an Hafer und Weizen. Lund Univ. Arsskrift, N. F. Afd. 21 Bd 5. Nr 2.

Opsahl, B. 1956. The discrimination of interactions and linkage in continuous variation. Biometrics 12: 415-432.

Pollak, E., Robinson, H. F., and Comstock, R. E. 1957. Inter-population hybrids in open-pollinated varieties of maize. The Am. Nat. 91: 387-391.

Powers, L. 1936. The nature of the interaction of genes affecting four quantitative characters in a cross between Hordeum deficiens and Hordeum vulgare. Genetics 21: 398-420.

_____. 1939. Studies on the nature of the interactions of the genes differentiating quantitative characters in a cross between Lycopersicon esulentum and L. pimpinellifolium. Jour. Genetics 39: 139-168.

_____. 1941. Inheritance of quantitative characters in crosses involving two species of Lycopersicon. Jour. Agr. Research 63: 149-174.

_____. 1942. The nature of the series of environmental variances and the estimation of the genetic variances and the geometric means in crosses involving species of Lycopersicon. Genetics 27: 561-575.

_____. 1955. Components of variance method and partitioning method of genetic analysis applied to weight per fruit of tomato hybrid and parental populations. U. S. Dept. Agr. Tech. Bul. 1131.

Rasmusson, J. 1934. A contribution to the theory of quantitative character inheritance. Hereditas 18: 245-261.

_____. 1935. Studies on the inheritance of quantitative characters in Pisum. I. Preliminary note on the genetics of time of flowering. Hereditas 20: 159-180.

Robinson, H. F., Mann, T. J., and Comstock, R. E. 1954. An analysis of quantitative variability in Nicotiana tabacum. Heredity 8: 365-376.

_____, Khalil, A., Comstock, R. E., and Cockerham, C. C. 1958. Joint interpretation of heterosis and genetic variances in two open-pollinated varieties of corn and their cross. Genetics 43: 868-877.

Rumbaugh, M. D. and Lonnquist, J. H. 1959. Inbreeding depression of diallel crosses of selected lines of corn. Agron. Jour. 51: 407-412.

Ryder, E. J. 1958. The effects of complementary epistasis on the inheritance of a quantitative character, seed size in lima beans. Agron. Jour. 50: 298-301.

Sentz, J. C., Robinson, H. F., and Comstock, R. E. 1954. Relation between heterozygosis and performance in maize. Agron. Jour. 46: 514-520.

Shull, G. H. 1914. Duplicate genes for capsule-form in Bursa bursa-pastoris. Zeitschr. fur Indukt. Abstamm. und Vererb. 12: 97-149.

Sinnott, E. W. 1937. The relation of gene to character in quantitative inheritance. Natl. Acad. Sci. Proc. 23: 224-227.

Smith, H. H. 1952. Fixing transgressive vigor in Nicotiana rustica. In Gowen, J. W., ed. Heterosis pp. 161-174. Ames, Iowa, Iowa State College Press.

Sprague, G. F. and Federer, W. T. 1951. A comparison of variance components in corn yield trials: II. Error, year x variety, location x variety, and variety components. Agron. Jour. 43: 535-541.

_____ and Tatum, L. A. 1942. General vs. specific combining ability in single crosses of corn. Jour. Amer. Soc. Agron. 34: 923-932.

Stringfield, G. H. 1950. Heterozygosis and hybrid vigor in maize. Agron. Jour. 42: 145-152.

Srb, A. M. and Owen, R. D. 1957. General genetics. San Francisco, Calif., W. H. Freeman and Company.

Wright, S. 1922. The effects of inbreeding and crossbreeding on guinea pigs. U. S. Dept. Agr. Bul. 1121.

_____. 1935. The analysis of variance and the correlation between relatives with respect to deviations from an optimum. Jour. Genetics 30: 243-256.

ACKNOWLEDGMENTS

The writer is grateful to Dr. L. H. Penny for his guidance and encouragement throughout the course of this investigation and the preparation of the manuscript. Appreciation is also expressed to Drs. O. Kempthorne and S. P. H. Mandel for their invaluable help in the statistical phases of this investigation and for their helpful suggestions in the preparation of the manuscript.

APPENDIX

Introduction

In this section is shown the method used in obtaining formulas for covariances of the type Cov(Cov 1, Cov 2), where Cov 1 = $\widehat{Cov}(X,X'/k,k-1;n,n')$ and Cov 2 is another such covariance. The method used in obtaining a formula for $Var[\widehat{Cov}(X,X'/3,2;3,4)]$ also is shown.

Models and Methods

To obtain Cov(Cov 1, Cov 2) use was made of the fact that

$$Cov(X,Y) = E(XY) - E(X)E(Y) \quad .$$

Therefore

$$Cov(Cov\ 1, Cov\ 2) = E(Cov\ 1, Cov\ 2) - E(Cov\ 1)E(Cov\ 2) \quad .$$

$Var[\widehat{Cov}(X,X'/3,2;3,4)]$ was obtained by using

$$Var(X) = E(X^2) - (EX)^2 \quad .$$

To work out the necessary expectations, the following approximate models were used. We can let

w_i = the observed value of the i^{th} F_2 plant
$i = 1, 2, \ldots, n$

x_{ij} = the observed value of the j^{th} F_3 plant in the i^{th} F_2 sub-population
$j = 1, 2, \ldots, r$

y_{ijk} = the observed value of the k^{th} F_4 plant in the j^{th} F_3 sub-population in the i^{th} F_2 sub-population
$k = 1, 2, \ldots, s$

Here it is assumed the w_i's are normally, independently distributed variables as are the x_{ij}'s and y_{ijk}'s.

Ignoring replications, the analysis of the F_3 and F_4 data would

appear as

Analysis of variance of F_3 data

Source	M.S.	E.M.S.
Between F_2 sub-populations	M_2'	$\sigma_b^2 + r\sigma_a^2$
Bet. F_3 plants in F_2 sub-pop.	M_1'	σ_b^2

$\sigma_a^2 = \mathrm{Cov}(X,X'/2,1;3,3)$

Analysis of variance of F_4 data

Source	M.S.	E.M.S.
Between F_2 sub-populations	M_3	$\sigma_1^2 + s\sigma_2^2 + sr\sigma_3^2$
Bet. F_3 sub-pop. in F_4 sub-pop.	M_2	$\sigma_1^2 + s\sigma_2^2$
Bet. F_4 plants in F_3 sub-pop. in F_2 sub-pop.	M_1	σ_1^2

$\sigma_2^2 = \mathrm{Cov}(X,X'/3,2;4,4)$

$\sigma_3^2 = \mathrm{Cov}(X,X'/2,1;4,4)$

Thus

$$\widehat{\mathrm{Cov}}(X,X'/2,1;3,3) = \frac{M_2' - M_1'}{r}$$

$$= \frac{1}{r}\left[\frac{r\sum_i (x_{i.} - x_{..})^2}{n-1} - \frac{\sum_{ij}(x_{ij} - x_{i.})^2}{n(r-1)}\right]$$

$$= \frac{\sum_i (x_{i.} - x_{..})^2}{n-1} - \frac{\sum_{ij}(x_{ij} - x_{i.})^2}{nr(r-1)} ,$$

$$\widehat{\mathrm{Cov}}(X,X'/3,2;4,4) = \frac{M_2 - M_1}{s}$$

$$= \frac{\sum_{ij}(y_{ij.} - y_{i..})^2}{n(r-1)} \quad \frac{\sum_{ijk}(y_{ijk} - y_{ij.})^2}{nrs(s-1)}$$

and

$$\widehat{Cov}(X,X'/2,1;4,4) = \frac{M_3 - M_2}{rs}$$

$$= \frac{\sum_i (y_{i..} - y_{...})^2}{n-1} - \frac{\sum_{ij} (y_{ij.} - y_{i..})^2}{nr(r-1)}$$

The other covariances are

$$\widehat{Cov}(X,X'/2,1;2,3) = \frac{\sum_i (w_i - w_.)(x_{i.} - x_{..})}{n-1}$$

$$\widehat{Cov}(X,X'/2,1;2,4) = \frac{\sum_i (w_i - w_.)(y_{i..} - y_{...})}{n-1}$$

$$\widehat{Cov}(X,X'/2,1;3,4) = \frac{\sum_i (x_{i.} - x_{..})(y_{i..} - y_{...})}{n-1}$$

$$\widehat{Cov}(X,X'/3,2;3,4) = \frac{\sum_{ij} (x_{ij} - x_{i.})(y_{ij.} - y_{i..})}{n(r-1)}$$

In the above formulas, estimates are denoted by hats. Dots in the subscripts of w, x, or y denote a mean of the observations taken over the dotted subscripts.

In the evaluation of the expectations, various combinations of F_2, F_3, and F_4 observations are involved. In order to simplify the calculations regression techniques were used. The principle used was to regress the observed value of the later generations on the observed values of the earlier generations.

If only F_3 and F_4 data were involved, y_{ijk} was regressed on x_{ij}. The model used was

$$y_{ijk} = \alpha_1 + B_1 X_{ij} + e_{ijk}$$

where e_{ijk} is NID(0, σ_e^2). Therefore

$$y_{ijk} - y_{ij.} = e_{ijk} - e_{ij.}$$

$$y_{ij.} - y_{i..} = B_1(x_{ij} - x_{i.}) + (e_{ij.} - e_{i..})$$

$$y_{i..} - y_{...} = B_1(x_{i.} - x_{..}) + (e_{i..} - e_{...})$$

x_{ij} has a structure which can be represented by

$$x_{ij} = u + a_i + b_{ij}$$

where u is a constant, a_i is the F_2 effect, and b_{ij} is the F_3 effect in the i^{th} F_2. From this model we obtain

$$x_{ij} - x_{i.} = b_{ij} - b_{i.}$$

$$x_{i.} - x_{..} = (a_i - a_.) + (b_{i.} - b_{..})$$

and then

$$y_{ijk} - y_{ij.} = e_{ijk} - e_{ij.}$$

$$y_{ij.} - y_{i..} = B_1(b_{ij} - b_{i.}) + (e_{ij.} - e_{i..})$$

$$y_{i..} - y_{...} = B_1(a_i - a_.) + B_1(b_{i.} - b_{..}) + (e_{i..} - e_{...})$$

If only F_2 and F_4 data were involved in Cov(Cov 1, Cov 2), then y_{ijk} was regressed on w_i and the model used was

$$y_{ijk} = \alpha_2 + B_2 w_i + f_{ij} + g_{ijk}$$

where f_{ij} and g_{ijk} are error terms and are normally and independently distributed with zero means and variances, σ_f^2 and σ_g^2, respectively. Using this model

$$y_{ijk} - y_{ij.} = g_{ijk} - g_{ij.}$$

$$y_{ij.} - y_{i.} = (f_{ij} - f_{i.}) + (g_{ij.} - g_{i..})$$

$$y_{i..} - y_{...} = B_2(w_i - w_.) + (f_{i.} - f_{..}) + (g_{i..} - g_{...})$$

If in Cov(Cov 1,Cov 2) only F_2 and F_3 data were involved, then x_{ij} was regressed on w_i. The model was

$$x_{ij} = \alpha_3 + B_3 w_i + h_{ij}$$

where h_{ij} is $NID(0, \sigma_h^2)$. Therefore

$$x_{ij} - x_{i.} = h_{ij} - h_{i.}$$

$$x_{i.} - x_{..} = B_3(w_i - w_.) + (h_{i.} - h_{..})$$

Where F_2, F_3, and F_4 data were involved in Cov(Cov 1,Cov 2) then y_{ijk} and x_{ij} were regressed on w_i. The models were the same as those already given.

To obtain the expectations of (Cov 1,Cov 2) and $\widehat{Cov}(X,X'/3,2;3,4)$ terms of the form $E(z_i - z_.)^r$ had to be evaluated. If $(z_i - z_.)^r$ is expanded and thn expectations taken, this will involve moments about the origin, i.e.

$$E(z_i^k) = u_k' \quad ,$$

where

$$k = 1, 2, \ldots, r,$$

and following standard notation

$$u_1' = u \ .$$

If the z_i's are random variables, then $E(z_i - z_.)^r$ is invariant with respect to the change of origin. We can change the origin from the original value to the mean of the distribution and then $E(z_i - z_.)^r$ will be expressed in terms of moments about the mean, i.e.

$$E(z_i - u)^k = u_k \ .$$

For the moments about the mean

$$u_1 = 0 \quad \text{and} \quad u_2 = \sigma^2 \ .$$

That $E(z_i - z_.)^r$ is invariant with respect to change of origin can be seen by the following reasoning. Given a random variable x_i, we define a new variable

$$y_i = x_i - u \quad ,$$

i.e. we shift the origin to the mean, u. Now

$$E(y_i)^k = E(x_i - u)^k = u_k \quad .$$

Since

$$y_. = \frac{y_1 + y_2 + \ldots + y_n}{n}$$

$$= x_. - u$$

then

$$y_i - y_. = (x_i - u) - (x_. - u)$$

$$= x_i - x_.$$

Now $E(x_i - x_.)^r$ will be a certain expression involving the u_k''s, and $E(y_i - y_.)^r$ will involve the u_k's. Since

$$x_i - x_. = y_i - y_. \quad ,$$

then $E(x_i - x_.)^r$ can be expressed in terms of moments about the origin, u_k's.

In the derivation of Cov(Cov 1,Cov 2) and $\mathrm{Var}[\widehat{\mathrm{Cov}}(X,X'/3,2;3,4)]$ use was made of the fact that if a random variable is normally distributed then

$$u_4 = 3u_2^2 = 3\sigma^4 \quad .$$

It will be noted that the variables w_i, x_{ij}, and y_{ijk} that enter into the formulas for Cov(Cov 1,Cov 2) and $\mathrm{Var}[\widehat{\mathrm{Cov}}(X,X'/3,2;3,4)]$, were assumed to be normally distributed random variables.

Derivation of Cov(Cov 1,Cov 2)

In this study 21 Cov(Cov 1,Cov 2)'s were evaluated. The method given by Kempthorne (1957) was used to obtain Cov[$\widehat{Cov}(X,X'/2,1;4,4)$, $\widehat{Cov}(X,X'/3,2;4,4)$]. Formulas for the other 20 Cov(Cov 1,Cov 2)'s were derived in the course of this investigation. To show the method used in obtaining these formulas, the derivation of the formula for Cov[$\widehat{Cov}(X,X'/2,1;3,3)$, $\widehat{Cov}(X,X'/2,1;2,3)$] is shown in detail. The results of the others are reported.

Let

$$\text{Cov 1} = \widehat{Cov}(X,X'/2,1;3,3) = \frac{\sum_i (x_{i.} - x_{..})^2}{n-1} - \frac{\sum_{ij} (x_{ij} - x_{i.})^2}{nr(r-1)}$$

$$\text{Cov 2} = \widehat{Cov}(X,X'/2,1;2,3) = \frac{\sum_i (w_i - w_.)(x_{i.} - x_{..})}{n-1} \quad .$$

We want

$$\text{Cov(Cov 1,Cov 2)} = E(\text{Cov 1,Cov 2}) - E(\text{Cov 1})\,E(\text{Cov 2}) \quad .$$

Now E(Cov 1,Cov 2) equals

$$\frac{1}{(n-1)^2} E\left[\sum_i (w_i - w_.)(x_{i.} - x_{..}) \sum_m (x_{m.} - x_{..})^2\right]$$

$$- \frac{1}{n(n-1)r(r-1)} E\left[\sum_i (w_i - w_.)(x_{i.} - x_{..}) \sum_{mj} (x_{mj} - x_{m.})^2\right]$$

$$= \frac{E(1)}{(n-1)^2} - \frac{E(2)}{n(n-1)r(r-1)}$$

$i = 1, 2, \ldots, n$

$m = 1, 2, \ldots, n$

$j = 1, 2, \ldots, r$

To obtain the above expectations x_{ij} is regressed on w_i. The model was given previously.

We find

$$E(1) = E\sum_i (w_i - w_.)(x_{i.} - x_{..})(x_{i.} - x_{..})^2$$

$$+ E\sum_{\substack{im \\ i\neq m}} (w_i - w_.)(x_{i.} - x_{..})(x_{m.} - x_{..})^2$$

$$= nE(w_i - w_.)(x_{i.} - x_{..})(x_{i.} - x_{..})^2$$

$$+ n(n-1)E(w_i - w_.)(x_{i.} - x_{..})(x_{m.} - x_{..})^2$$

$$= n\,E(A) + n(n-1)E(B) \quad .$$

Substitute

$$x_{i.} - x_{..} = B_3(w_i - w_.) + (h_{i.} - h_{..})$$

into the above expressions. This gives

$$E(A) = B_3^3\, E(w_i - w_.)^4 + 3B_3E(w_i - w_.)^2(h_{i.} - h_{..})^2 \quad .$$

$E(w_i - w_.)^4$ is obtained by expanding $(w_i - w_.)^4$ and then taking expectations of the individual terms. Detailed steps in the derivation of this expectation will be shown to acquaint the reader with the method of obtaining expectations of the type $E(z_i - z_.)^r$.

Expanding $(w_i - w_.)^4$ we have

$$E(w_i - w_.)^4 = E(w_i^4 - 4w_i^3w_. + 6w_i^2w_.^2 - 4w_iw_.^3 + w_.^4) \quad .$$

Evaluating the individual terms of the above expectation, we have

$$E(w_i^4) = u_4'$$

$$E(w_i^3 w_\cdot) = E(w_i^3)\left(\frac{w_1 + \cdots + w_i + \cdots + w_n}{n}\right)$$

$$= \frac{1}{n} E(w_1 w_i^3 + \cdots + w_i^4 + \cdots + w_i^3 w_n)$$

$$= \frac{u_4' + (n-1)uu_3'}{n}$$

$$E(w_i^2 w_\cdot^2) = \frac{1}{n^2} E(w_i^2)\left(\sum_i w_i^2 + \sum_{i \neq j} w_i w_j\right)$$

$$= \frac{u_4' + (n-1)(u_2')^2 + 2(n-1)uu_3' + (n-1)(n-2)u^2 u_2'}{n^2}$$

$$E(w_i w_\cdot^3) = Ew_i \left[\frac{\sum w_i}{n}\right]^3$$

$$= \frac{1}{n^3}\left[u_4' + 4(n-1)uu_3' + 3(n-1)(u_2')^2 + 6(n-1)(n-2)u^2 u_2' + (n-1)(n-2)(n-3)u^4\right]$$

$$E(w_\cdot^4) = \frac{1}{n^4} E\left(\sum_i w_i^4 + 4\sum_{i \neq j} w_i^3 w_j + 3\sum_{i \neq j} w_i^2 w_j^2 + 12\sum_{i \neq j \neq k} w_i^2 w_j w_k + \sum_{i \neq j \neq k \neq m} w_i w_j w_k w_m\right)$$

$$= \frac{1}{n^3}\left[u_4' + 4(n-1)uu_3' + 3(n-1)(u_2')^2 + 6(n-1)(n-2)u^2 u_2' + (n-1)(n-2)(n-3)u^4\right]$$

$$= E(w_i w_\cdot^3) \quad .$$

Combining terms of $E(w_i - w_\cdot)^4$ we will obtain an expression involving moments about the origin. But $E(w_i - w_\cdot)^4$ is invariant with respect to

the change of the origin so it can be expressed in terms of moments about the mean. This can be done by taking $E(w_i - w_.)^4$, expressed in terms of moments about the origin, and setting u and u_3' equal to zero, u_2' equal to σ_w^2, and u_4' equal to $3\sigma_w^4$. Doing this we obtain

$$E(w_i - w_.)^4 = \frac{3(n-1)^2}{n^2} \sigma_w^4$$

In obtaining expectations like $E(w_i - w_.)^4$ it is possible to obtain the correct result by expressing the individual terms in moments about the mean instead of waiting to do this until after the individual terms are combined. This will result in considerable simplification. But in doing this we should realize that

$$E(w_i^k) \neq u_k$$

and it is only when the individual terms are combined can it be shown mathematically that $E(w_i - w_.)^4$ is invariant which respect to the change of the origin.

To evaluate $E(w_i - w_.)^2 (h_{i.} - h_{..})$, use is made of the principle that if two variables, X and Y, are independent, then

$$E(XY) = E(X)E(Y) \quad .$$

Therefore

$$E(w_i - w_.)^2(h_{i.} - h_{..})^2 = E(w_i - w_.)^2 E(h_{i.} - h_{..})^2 \quad .$$

Evaluating these terms by the same method as used for $E(w_i - w_.)^4$ we obtain

$$E(w_i - w_.)^2 = \frac{n-1}{n} \sigma_w^2 \quad ,$$

and

$$E(h_i - h_{..})^2 = \frac{n-1}{nr}\sigma_h^2 \quad .$$

Combining all the terms for E(A), gives

$$E(A) = \frac{3(n-1)^2}{n^2 r} B_3 \sigma_w^2(\sigma_h^2 + rB_3^2\sigma_w^2) \quad .$$

Substituting

$$x_{i.} - x_{..} = B_3(w_i - w_.) + (h_{i.} - h_{..})$$

into E(B), we have

$$E(B) = B_3^3 E(w_i - w_.)^2(w_m - w_.)^2 + B_3 E(w_i - w_.)^2(h_{m.} - h_{..})^2$$
$$+ 2B_3 E(w_i - w_.)(w_m - w_.)(h_{i.} - h_{..})(h_{m.} - h_{..})$$

$$= B_3^3 \frac{n^2-2n+3}{n^2}\sigma_w^4 + B_3 \frac{(n-1)^2}{n^2 r}\sigma_w^2\sigma_h^2 + 2B_3 \frac{1}{n^2 r}\sigma_w^2\sigma_h^2$$

$$= \frac{n^2-2n+3}{n^2 r} B_3 \sigma_w^2(\sigma_h^2 + rB_3^2\sigma_w^2) \quad .$$

Combining terms

$$E(1) = nE(A) + n(n-1)E(B)$$

$$= \frac{(n-1)(n+1)}{r} B_3 \sigma_w^2(\sigma_h^2 + rB_3^2\sigma_w^2) \quad .$$

We next need E(2). This is

$$E(2) = E\left[\sum_i (w_i - w_.)(x_{i.} - x_{..}) \sum_{mj} (x_{mj} - x_{m.})^2\right]$$

$$= E\sum_i (w_i - w_.)(x_{i.} - x_{..})(x_{ij} - x_{i.})^2$$
$$+ E\sum_{i \neq m} (w_i - w_.)(x_{i.} - x_{..})(x_{mj} - x_{m.})^2$$

$$= nrE(w_i - w_.)(x_{i.} - x_{..})(x_{ij} - x_{i.})^2$$
$$+ nr(n-1)E(w_i - w_.)(x_{i.} - x_{..})(x_{mj} - x_{m.})^2$$

$$= nrE(C) + nr(n-1)E(D) \quad .$$

Substituting

$$(x_{i.} - x_{..}) = B_3(w_i - w_.) + (h_{i.} - h_{..})$$

and

$$x_{ij} - x_{i.} = h_{ij} - h_{i.}$$

into E(C) and E(D) we obtain

$$E(C) = B_3E(w_i - w_.)^2(h_{ij} - h_{i.})^2$$

$$= \frac{(n-1)(r-1)}{nr} \; B_3 \sigma_w^2 \sigma_h^2$$

and

$$E(D) = B_3E(w_i - w_.)^2(h_{mj} - h_{m.})^2$$

$$= E(C) \quad .$$

Therefore

$$E(2) = n^2rE(C)$$

$$= n(n-1)(r-1)B_3 \sigma_w^2 \sigma_h^2 \quad .$$

We want

$$E(\text{Cov } 1, \text{Cov } 2) = \frac{E(1)}{(n-1)^2} - \frac{E(2)}{n(n-1)r(r-1)} \quad .$$

Combining terms this is

$$E(\text{Cov } 1, \text{Cov } 2) = \frac{n+1}{n-1} B_3^3 \sigma_w^4 + \frac{2}{(n-1)r} B_3 \sigma_w^2 \sigma_h^2 \quad .$$

To obtain Cov(Cov 1,Cov 2) the expectations of Cov 1 and Cov 2 are needed. Now

$$E(\text{Cov } 1) = E[\widehat{\text{Cov}}(X,X'/2,1;3,3)]$$

$$= E\left[\frac{\sum_i (x_{i.} - x_{..})^2}{n-1}\right] - E\left[\frac{\sum_{ij} (x_{ij} - x_{i.})^2}{nr(r-1)}\right]$$

$$= \frac{n}{n-1} E(x_{i.} - x_{..})^2 - \frac{nr}{nr(r-1)} E(x_{ij} - x_{i.})^2$$

$$= \frac{n}{n-1} E[B_3(w_i - w_.) + (h_{i.} - h_{..})] - \frac{1}{r-1} E(h_{ij} - h_{i.})^2$$

$$= B_3^2 \sigma_w^2$$

and

$$E(\text{Cov } 2) = E[\text{Cov}(X, X'/2,1;2,3)]$$

$$= E\left[\frac{\sum_i (w_i - w_.)(x_{i.} - x_{..})}{n-1}\right]$$

$$= \frac{n}{n-1} E[B_3(w_i - w_.)^2 + (w_i - w_.)(h_{i.} - h_{..})]$$

$$= B_3 \sigma_w^2$$

Therefore

$$E(\text{Cov } 1)E(\text{Cov } 2) = B_3^3 \sigma_w^4$$

Using the above results Cov(Cov 1,Cov 2) now can be obtained. The result is

$$\text{Cov}(\text{Cov } 1, \text{Cov } 2) = E(\text{Cov } 1, \text{Cov } 2) - E(\text{Cov } 1)E(\text{Cov } 2)$$

$$= \frac{2}{(n-1)r} B_3 \sigma_w^2 (\sigma_h^2 + rB_3^2 \sigma_w^2) \quad .$$

Formulas for the other 19 Cov(Cov 1,Cov 2)'s were obtained by the same methods as used in the derivation of $\text{Cov}[\widehat{\text{Cov}}(X,X'/2,1;3,3), \widehat{\text{Cov}}(X,X'/2,1;2,3)]$. Formulas for the 20 Cov(Cov 1,Cov 2)'s derived in this study are shown in table 12.

Table 12. Formulas for Cov(Cov 1,Cov 2)'s

Cov 1	Cov 2	Cov(Cov 1,Cov 2)
2,1;3,3	2,1;2,3	$\frac{2}{(n-1)r} B_3 \sigma_w^2(\sigma_h^2 + rB_3^2\sigma_w^2)$
	2,1;4,4	$\frac{2}{(n-1)r^2} B_1^2(\sigma_b^2 + r\sigma_a^2)^2$
	3,2;4,4	$-\frac{2}{nr(r-1)} B_1^2\sigma_b^4$
	2,1;2,4	$\frac{2}{n-1} B_2B_3^2\sigma_w^4$
	2,1;3,4	$\frac{2}{(n-1)r^2} B_1(\sigma_b^2 + r\sigma_a^2)^2$
	3,2;3,4	$-\frac{2}{nr(r-1)} B_1\sigma_b^4$
2,1;2,3	2,1;4,4	$\frac{2}{n-1} B_2^2B_3\sigma_w^4$
	3,2;4,4	0
	2,1;2,4	$\frac{2}{n-1} B_2B_3\sigma_w^4$
	2,1;3,4	$\frac{1}{(n-1)r} B_2\sigma_w^2(\sigma_n^2 + 2rB_3^2\sigma_w^2)$
	3,2;3,4	0
2,1;4,4	2,1;2,4	$\frac{2}{(n-1)rs} B_2\sigma_w^2(\sigma_g^2 + s\sigma_f^2 + rsB_2^2\sigma_w^2)$
	2,1;3,4	$\frac{2}{(n-1)r^2s} B_1(\sigma_b^2\sigma_e^2 + r\sigma_a^2\sigma_e^2 + sB_1^2\sigma_b^4 + 2rsB_1^2\sigma_a^2\sigma_b^2 + r^2sB_1^2\sigma_a^4)$
	3,2;3,4	$-\frac{2}{nr(r-1)s} B_1\sigma_b^2(\sigma_e^2 + sB_1^2\sigma_b^2)$
3,2;4,4	2,1;2,4	0
	2,1;3,4	0

Table 12. (Continued)

Cov 1	Cov 2	Cov(Cov 1, Cov 2)
3,2;4,4	3,2;3,4	$\frac{2}{n(r-1)s} B_1 \sigma_b^2(\sigma_e^2 + sB_1^2\sigma_b^2)$
2,1;2,4	2,1;3,4	$\frac{1}{(n-1)rs} B_3 \sigma_w^2(\sigma_g^2 + s\sigma_f^2 + 2rsB_2^2\sigma_w^2)$
	3,2;3,4	0
2,1;3,4	3,2;3,4	0

By use of the methods described here

$$\begin{aligned}\mathrm{Cov}[\widehat{\mathrm{Cov}}(X,X'/2,1;2,3),\widehat{\mathrm{Cov}}(X,X'/3,2;3,4)] &= \mathrm{Cov}[\widehat{\mathrm{Cov}}(X,X'/2,1;2,4),\widehat{\mathrm{Cov}}(X,X'/3,2;3,4)] \\ &= 0\end{aligned}$$

was obtained. In the derivation of these, it was found that the approximate models used were not exact enough to give a reasonable expectation of $\widehat{\mathrm{Cov}}(X,X'/3,2;3,4)$.

Because the estimator of $\mathrm{Cov}(X,X'/3,2;3,4)$ is unbiased, then by definition

$$E[\widehat{\mathrm{Cov}}(X,X'/3,2;3,4)] = \mathrm{Cov}(X,X'/3,2;3,4) \quad .$$

In the above two covariances x_{ij} and y_{ijk} were regressed on w_i. When $\widehat{\mathrm{Cov}}(X,X'/3,2;3,4)$ was expressed in terms of such regressions, $E[\widehat{\mathrm{Cov}}(X,X'/3,2;3,4)]$ equalled zero. This expectation obviously is not correct. An examination of the factors going into these two covariances led to the conclusion that even under exact models where $E[\widehat{\mathrm{Cov}}(X,X'/3,2;3,4)]$ would be more realistic, these covariances would be zero. Therefore, answers obtained by use of the approximate models were assumed

correct.

Estimators of Cov(Cov 1,Cov 2)'s

In order to evaluate the 20 Cov(Cov 1,Cov 2)'s given in table 12 estimators of these covariances had to be derived. To acquaint the reader with the method used to derive these estimators the derivation of the estimator for $\mathrm{Cov}[\widehat{\mathrm{Cov}}(X,X'/2,1;3,3),\widehat{\mathrm{Cov}}(X,X'/2,1;2,3)]$ is shown below.

$\mathrm{Cov}[\widehat{\mathrm{Cov}}(X,X'/2,1;3,3),\widehat{\mathrm{Cov}}(X,X'/2,1;2,3)]$ is equal to

$$\frac{2}{(n-1)r} B_3 \sigma_w^2 (\sigma_h^2 + rB_3^2 \sigma_w^2)$$

As shown previously the expectation of $\widehat{\mathrm{Cov}}(X,X'/2,1;2,3)$ is $B_3\sigma_w^2$. An unbiased estimate of the variance of the F_3 means is

$$\frac{\sum_i (x_{i.} - x_{..})^2}{n-1}$$

The expectation of this estimate is

$$E\left[\frac{\sum_i (x_{i.} - x_{..})^2}{n-1}\right] = \frac{n}{n-1} E[B_3(w_i - w_.) + (h_{i.} - h_{..})]^2$$

$$= B_3^2\sigma_w^2 + \frac{1}{r}\sigma_h^2$$

Therefore

$$r\sigma_{x_{i.}}^2 = \sigma_h^2 + rB_3^2\sigma_w^2$$

An unbiased estimate of $\mathrm{Cov}[\widehat{\mathrm{Cov}}(X,X'/2,1;3,3),\widehat{\mathrm{Cov}}(X,X'/2,1;2,3)]$ would be

$$\frac{2}{n-1}[\widehat{\mathrm{Cov}}(X,X'/2,1;2,3)]\,\widehat{\sigma}_{x_{i.}}^2$$

The estimators used to evaluate the Cov(Cov 1,Cov 2)'s given in table 12 are shown in table 13.

Table 13. Estimators of Cov(Cov 1,Cov 2)'s

Cov 1	Cov 2	Estimator of Cov(Cov 1,Cov 2)
2,1;3,3	2,1;2,3	$\frac{2}{n-1}[\hat{Cov}(X,X'/2,1;2,3)]\hat{\sigma}^2_{x_i.}$
	2,1;4,4	$\frac{2}{n-1}[\hat{Cov}(X,X'/2,1;3,4)]^2$
	3,2;4,4	$-\frac{2}{nr(r-1)}[\hat{Cov}(X,X'/3,2;3,4)]^2$
	2,1;2,4	$\frac{2}{n-1}[\hat{Cov}(X,X'/2,1;3,3)][\hat{Cov}(X,X'/2,1;2,4)]$
	2,1;3,4	$\frac{2}{n-1}\frac{[\hat{Cov}(X,X'/2,1;3,4)]^2[\hat{Cov}(X,X'/3,2;3,4)]}{\hat{Cov}(X,X'/3,2;4,4)}$
	3,2;3,4	$-\frac{2}{nr(r-1)}\frac{[\hat{Cov}(X,X'/3,2;3,4)]^3}{\hat{Cov}(X,X'/3,2;4,4)}$
2,1;2,3	2,1;4,4	$\frac{2}{n-1}[\hat{Cov}(X,X'/2,1;2,3)][\hat{Cov}(X,X'/2,1;4,4)]$
	3,2;4,4	0
	2,1;2,4	$\frac{2}{n-1}[\hat{Cov}(X,X'/2,1;2,3)][\hat{Cov}(X,X'/2,1;2,4)]$
	2,1;3,4	$\frac{1}{n-1}[\hat{Cov}(X,X'/2,1;2,4)]\hat{\sigma}^2_{x_i.}$ $+ [\hat{Cov}(X,X'/2,1;2,3)][\hat{Cov}(X,X'/2,1;3,4)]$
2,1;2,3	3,2;3,4	0
2,1;4,4	2,1;2,4	$\frac{2}{n-1}[\hat{Cov}(X,X'/2,1;2,4)]\hat{\sigma}^2_{y_i..}$
	2,1;3,4	$\frac{2}{n-1}[\hat{Cov}(X,X'/2,1;3,4)]\hat{\sigma}^2_{y_i..}$
	3,2;3,4	$-\frac{2}{nr(r-1)}[\hat{Cov}(X,X'/3,2;3,4)]\frac{\sum_{ij}(y_{ij.}-y_{i..})^2}{n(r-1)}$

Table 13. (Continued)

Cov 1	Cov 2	Estimator of Cov(Cov 1,Cov 2)
3,2;4,4	2,1;2,4	0
	2,1;3,4	0
	3,2;3,4	$\frac{2}{n(r-1)}\,[\widehat{Cov}(X,X'/3,2;3,4)]\,\frac{\sum_{ij}(y_{ij.}-y_{i..})^2}{n(r-1)}$
2,1;2,4	2,1;3,4	$\frac{1}{n-1}\,[\widehat{Cov}(X,X'/2,1;2,3)]\hat{\sigma}^2_{y_{i..}}$ + $[\widehat{Cov}(X,X'/2,1;2,4)][\widehat{Cov}(X,X'/2,1;3,4)]$
	3,2;3,4	0
2,1;3,4	3,2;3,4	0

Variance of $\widehat{Cov}(X,X'/3,2;3,4)$

The same procedure that was used to obtain the Cov(Cov 1,Cov 2)'s was employed in obtaining the variance of $\widehat{Cov}(X,X'/3,2;3,4)$. Here

$$Var(X) = E(X^2) - (EX)^2$$

Using this relationship the following formula was obtained

$$Var[\widehat{Cov}(X,X'/3,2;3,4)] = \frac{1}{n(r-1)s}\,\sigma^2_b(\sigma^2_e + 2sB^2_1\sigma^2_b)$$

An unbiased estimate of this variance is

$$\frac{1}{n(r-1)}\,[\widehat{Cov}(X,X'/3,2;3,4)]^2\left[\frac{1}{\widehat{Cov}(X,X'/3,2;4,4)}\,\frac{\sum_{ij}(y_{ij.}-y_{i..})^2}{n(r-1)} + 1\right].$$